

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 09:48:23 ; Search time 36.5 Seconds
(without alignments)
77.799 Million cell updates/sec

Title: US-09-980-064-1
Perfect score: 55
Sequence: 1 AQIFNKPYW 9

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315510202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	44	12	Q84308 human papill
2	55	100.0	44	12	Q84307 human papill
3	55	100.0	208	12	Q84306 human papill
4	55	100.0	211	12	Q84305 human papill
5	55	100.0	221	12	Q84304 human papill
6	55	100.0	419	12	Q84303 human papill
7	55	100.0	494	12	Q84302 human papill
8	55	100.0	494	12	Q84301 human papill
9	55	100.0	494	12	Q84300 human papill
10	55	100.0	494	12	Q84299 human papill
11	55	100.0	494	12	Q84298 human papill
12	55	100.0	494	12	Q84297 human papill
13	55	100.0	494	12	Q84296 human papill
14	55	100.0	494	12	Q84295 human papill
15	55	100.0	494	12	Q84294 human papill
16	55	100.0	494	12	Q84293 human papill

17	55	100.0	494	12	Q80997 human papill
18	55	100.0	494	12	Q81008 human papill
19	55	100.0	494	12	Q80999 human papill
20	55	100.0	494	12	Q81009 human papill
21	55	100.0	494	12	Q81012 human papill
22	55	100.0	494	12	Q81001 human papill
23	55	100.0	494	12	Q90204 human papill
24	55	100.0	494	12	Q81000 human papill
25	55	100.0	495	12	Q12343 human papill
26	55	100.0	504	12	Q81017 human papill
27	55	100.0	505	12	Q9WH57 human papill
28	55	100.0	505	12	Q9WLO5 human papill
29	55	100.0	505	12	Q9WH56 human papill
30	55	100.0	505	12	Q81015 human papill
31	55	100.0	505	12	Q9WPH4 human papill
32	55	100.0	505	12	Q92282 human papill
33	55	100.0	505	12	Q9WH55 human papill
34	55	100.0	505	12	Q9WLO6 human papill
35	55	100.0	531	12	Q00530 human papill
36	55	100.0	531	12	Q8B5N9 human papill
37	55	100.0	531	12	Q8B5N5 human papill
38	55	100.0	531	12	Q8WML8 human papill
39	55	100.0	565	12	Q06951 human papill
40	53	96.4	44	12	Q84304 human papill
41	53	96.4	44	12	Q84315 human papill
42	53	96.4	199	12	Q84301 human papill
43	53	96.4	485	12	Q9Q2T7 human papill
44	53	96.4	500	12	Q9WBR8 human papill
45	53	96.4	500	12	Q9W9C6 human papill

ALIGNMENTS

RESULT 1
Q84308 PRELIMINARY; PRT; 44 AA.
AC Q84308:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE L1 (Fragment).
GN L1.
OS Human papillomavirus type 34.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10613;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92407963; PubMed=1326639;
RA Chan S.-Y., Bernard H.-U., Ong C.-K., Chan S.-P., Birgit H.,
RA Delius H.,
RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and
RT variants: A showcase for the molecular evolution of DNA viruses.";
RL J. Virol. 66:5714-5725(1992).
DR EMBL; M96292; AAA47031.1; -;
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR002210; PV capsid L1.
DR InterPro; IPR008975; Viral cap coat.
DR Pfam; PF00500; late protein L1; 1.
DR Pfam; PF00544; PV capsid L1; 1.
FT NON_TER
SQ SEQUENCE 44 AA; 5075 MW; F7031326D89E1F2C CRC64;

Query Match 100.0%; Score 55; DB 12; Length 44;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
|||
DB 3 AQIFNKPYW 11

```

RESULT 2
Q84307 PRELIMINARY; PRT; 44 AA.
AC Q84307;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE L1 (Fragment).
GN L1.
OS Human papillomavirus type 32.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCB1_TaxID=10612;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92407963; PubMed=1326639;
RA Chan S.-Y., Bernard H.-U., Ong C.-K., Chan S.-P., Birgit H.,
RA Dellus H.,
RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and
RT variants: A showcase for the molecular evolution of DNA viruses.";
RL J. Virol. 66:5714-5725 (1992).
DR EMBL; M96291; AAA47030.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
SQ SEQUENCE 44 AA; 4988 MW; 1C396FA3DA4E048D CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 44;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPW 9
DB 3 AQIFNKPW 11

RESULT 3
Q8JMU6 PRELIMINARY; PRT; 208 AA.
AC Q8JMU6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCB1_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bsb-102;
RA Cruz M.R., Cerqueira D.M., Camara G.N.L., Silva E.O., Carvalho L.G.S.,
RA Martins C.R.F.;
RT "Human papillomavirus type 16 variants in Central Brazil.";
RT Submitted (APR-2002) to the EMBL/GenBank/DBD databases.
DR EMBL; AY098925; AAM29173.1; -
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
SQ SEQUENCE 208 AA; 23455 MW; 936AADD7ADE483 CRC64;

```

```

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 208;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPW 9
DB 46 AQIFNKPW 54

RESULT 4
Q8JMU9 PRELIMINARY; PRT; 211 AA.
AC Q8JMU9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCB1_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bsb-178;
RA Cruz M.R., Cerqueira D.M., Camara G.N.L., Silva E.O., Carvalho L.G.S.,
RA Martins C.R.F.;
RT "Human papillomavirus type 16 variants in Central Brazil.";
RT Submitted (APR-2002) to the EMBL/GenBank/DBD databases.
DR EMBL; AY098921; AAM29169.1; -
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
SQ SEQUENCE 211 AA; 23840 MW; 28134419632A1E9D CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 211;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPW 9
DB 49 AQIFNKPW 57

RESULT 5
Q8JMU5 PRELIMINARY; PRT; 221 AA.
AC Q8JMU5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCB1_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bsb-14;
RA Cruz M.R., Cerqueira D.M., Camara G.N.L., Silva E.O., Carvalho L.G.S.,
RA Martins C.R.F.;
RT "Human papillomavirus type 16 variants in Central Brazil.";
RT Submitted (APR-2002) to the EMBL/GenBank/DBD databases.
DR EMBL; AY098926; AAM29174.1; -
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.

```

```

DR InterPro; IPR002210; PV capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 221
SQ SEQUENCE 221 AA; 25004 MW; 96C9F3B4B4BD218 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 221;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AQIFNKPYW 9
DB 52 AQIFNKPYW 60

RESULT 6
O8JQM4 PRELIMINARY; PRT; 419 AA.
AC O8JQM4;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE l1 protein (Fragment).
GN l1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang J., Chen R.;
RT "Cloning and sequence analysis of Human papillomavirus type 16 L1 gene
RT in China.";
RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF512011; AAM74159.1; -.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR002210; PV capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 419
SQ SEQUENCE 419 AA; 46880 MW; 82C99A0D5F584560 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 419;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AQIFNKPYW 9
DB 329 AQIFNKPYW 337

RESULT 7
O81010 PRELIMINARY; PRT; 494 AA.
ID O81010;
AC O81010;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE late major capsid protein l1 (Fragment).
GN l1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=OR 7632;
RC STRAIN=OR 7632;
SQ SEQUENCE 494 AA; 55090 MW; 36335F09ECAFF3CE CRC64;

```

```

RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
RT populations characterized by nucleotide sequence analysis of the B6,
RT l2, and l1 coding segments.";
RL J. Virol. 69:7743-7753(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Farmer A.D.;
RC STRAIN=OR 7632;
RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U34174; AAA91736.1; -.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 494
FT NON_TER 494
SQ SEQUENCE 494 AA; 55094 MW; ECB31B09E8EFB4E3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 494;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AQIFNKPYW 9
DB 295 AQIFNKPYW 303

RESULT 8
O81002 PRELIMINARY; PRT; 494 AA.
ID O81002;
AC O81002;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE late major capsid protein l1 (Fragment).
GN l1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=OR 3136;
RC STRAIN=OR 3136;
RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U34174; AAA91721.1; -.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 494
FT NON_TER 494
SQ SEQUENCE 494 AA; 55090 MW; 36335F09ECAFF3CE CRC64;

```

Query Match	100.0%	Score 55;	DB 12;	Length 494;
Best Local Similarity	100.0%;	Pred. No. 0.052;		
Matches	9;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;
Q7	1	AOIEFKPYW	9	
Db	295	AOIEFKPYW	303	

RESULT	9
ID	081003
PRELIMINARY:	PRT; 494 AA.
DT	01-NOV-1986 (TrEMBLrel. 01, Created)
DT	01-NOV-1986 (TrEMBLrel. 01, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Late major capsid protein LI (fragment).
GN	LI.
OS	Human papillomavirus.
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC	Papillomavirus.
OX	NCBI_TaxID=10566;
OX	LI
RA	SEQUENCE FROM N.A.
RP	STRAIN=OR 3759;
RC	
RX	MEDLINE=66079021; PubMed=7494284;
RA	Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA	Johnson S.A.;
RT	"Human papillomavirus type 16 variant lineages in United States
RT	populations characterized by nucleotide sequence analysis of the E6,
RT	L2, and L1 coding segments.";
RT	J. Virol. 69:7743-7753(1995).
TL	

RP SEQUENCE FROM N.A.
RC STRAIN=OR 3759;
RA Farmer A.D.;
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases
DR EMBL, U34176; AAA81723.1; -;
DR GO; GO:0019028; C:vital capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV capsid L1.
DR InterPro; IPR008975; Viral capsid coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsIDL1.
DR Prodom; PDO00544; PV capsid_L1; 1.
FT NON_TER 1 494
FT NON_TER 1 494
SQ SEQUENCE 494 AA; 5510 MW; 0BF173B2EB831233 CRC64;

Query Match	100.0%;	Score 55;	DB 152;	Length 494;
Best Local Similarity	100.0%;	Pred. No. 0.052;		
Matches	9;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;
QY	1	AOIFENKPYW	9	
db	295	AOIFENKPYW	303	

RESULT	10
Q81006	
ID	Q81006
AC	Q81006
DT	01-NOV-1996 (TREMBlrel. 01, Created)
PT	01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT	01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE	Late major capsid protein L1 (Fragment).
GN	L1.
OS	Human papillomavirus.
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae
CC	Papillomavirus.
OX	NCBI_TaxID=10566;
RN	[1]
RP	SEQUENCE FROM N.A.

RC STRAIN=OR 6106;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildeheim A.,
RA Jensen S.A.;
RT "human papillomavirus type 16 variant lineages in United States
RT populations characterized by nucleotide sequence analysis of the E6
RT L2, and L1 coding segments,"
EL J. Virol. 69:7743-7753 (1995).
EN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=OR 6106;
 RA Farmer A.D.;
 RL Submitted (Aug-1995) to the EMBL/genbank/DBDJ databases
 DR EMBL; U04183; AAA81730.1;
 DR CG; GO:0019028; C:vital capsid, IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPRO02210; PV capsid.L1.
 DR InterPro; IPRO08975; Viral cap.coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsid.L1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 FT NON_TER 494 1 494
 FT TER 1 1
 SQ SEQUENCE 494 AA; 5500 MW; EBC46B09E8EAB1E3 CRC64;

Query Match	100.0%;	Score 55;	DB 12;	length 494;
Best Local Similarity	100.0%;	Pred. No. 0.052;		
Matches	9;	Conservative 0;	Mismatches 0;	Indels 0;
	1 AQTENKRYW	9		
Db	295 AQTENKRYW	303		

RESULT 11	ID	Q81011	PRELIMINARY;	PRT;	494 AA.
AC	Q81011;				
DT	01-NOV-1996 (TrEMBLrel. 01, Created)				
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)				
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)				
DE	Late major capsid protein L1 (Fragment).				
LN	L1.				
OS	Human papillomavirus.				
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;				
OC	Papillomavirus.				
OX	NCBI_TaxID=10566;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=OR 8160;				
RX	MEDLINE=56079021; PubMed=7494284;				
RA	Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,				
RA	Jeison S.A.;				
RT	"Human papillomavirus type 16 variant lineages in United States				
RT	populations characterized by nucleotide sequence analysis of the E6				
RT	L2, and L1 coding segments.;"				
RL	J. Virol. 69:7743-7753 (1995).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=OR 8160;				
RA	Farmer A.D.;				
RL	Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; U34190; AAA91737.1; '-'				
DR	GO; GO:0019028; C:Viral capsid; IEA.				
DR	GO; GO:0005198; F:structural molecule activity; IEA.				
DR	InterPro; IPR002210; PV_capsid_L1				
DR	InterPro; IPR008975; Viral_cap.coat.				
DR	Pfam; PF00500; Late_protein_L1; 1.				
DR	PRINTS; PR00865; HPVcapsidL1.				
DR	Prodom; PD000544; PV_capsid_L1; 1.				
FT	NON_TER	1	494		
FT	NON_TER	1	494		
QO	SEQUENCE	494 AA;	55063 MW;	735F7C2B86CAC58F	CRC64;

Query Match 100.0%; Score 55; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AOFNKPYW 9
Db 295 AOFNKPYW 303

RESULT 12

081005 PRELIMINARY; PRT; 494 AA.
AC 081005;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 5110;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the E6,
L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 5110;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the E6,
L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753(1995).
RN [2]

Query Match 100.0%; Score 55; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AOFNKPYW 9
Db 295 AOFNKPYW 303

RESULT 13

080998 PRELIMINARY; PRT; 494 AA.
AC 080998;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=NM T197;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the E6,
L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NM T197;
RA Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RL EMBL; U34177; AAA91724.1; -;
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR002210; PV capsid L1.
DR InterPro; IPR008975; Viral cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 494 AA; 55021 MW; E446FD38994FB3DD CRC64;

Query Match 100.0%; Score 55; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AOFNKPYW 9
Db 295 AOFNKPYW 303

RESULT 14

081004 PRELIMINARY; PRT; 494 AA.
AC 081004;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 4541;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the E6,
L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 4541;
RA Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RL EMBL; U34177; AAA91724.1; -;
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR002210; PV capsid L1.
DR InterPro; IPR008975; Viral cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 494 AA; 55021 MW; E446FD38994FB3DD CRC64;

SEQ SEQUENCE 494 AA; 55063 MW; 735F7C2B86CAC58F CRC64;

Query Match 100.0%; Score 55; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
DB 295 AQIFNKPYW 303

RESULT 15

Q81007 PRELIMINARY; PRT; 494 AA.

AC Q81007; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).

GN Human papillomavirus.
OS Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=10566;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=OR 6170;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
Jenson S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the E6,
L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753(1995).

[2]
RN SEQUENCE FROM N.A.
RP STRAIN=OR 6170;
RC Farmer A.D.;
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL; U34164; AAA91731.1; -;
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.

FT NON_TER 1 494
FT SEQUENCE 494 AA; 55051 MW; BD92AC39CB68D2CE CRC64;
SQ

Query Match 100.0%; Score 55; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
DB 295 AQIFNKPYW 303

RESULT 16

Q80996 PRELIMINARY; PRT; 494 AA.

AC Q80996; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).

GN Human papillomavirus.
OS Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=10566;
RN [1]
RP SEQUENCE FROM N.A.

RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=NM 4094;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
Jenson S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the E6,
L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753(1995).

[2]
RN SEQUENCE FROM N.A.
RP STRAIN=NM 4094;
RC Farmer A.D.;

RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U34165; AAA91712.1; -;
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.

FT NON_TER 1 494
FT SEQUENCE 494 AA; 55063 MW; 735F7C2B86CAC58F CRC64;
SQ

Query Match 100.0%; Score 55; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
DB 295 AQIFNKPYW 303

RESULT 17

Q80997 PRELIMINARY; PRT; 494 AA.

AC Q80997; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).

GN Human papillomavirus.
OS Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=10566;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=NM 9999;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
Jenson S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the E6,
L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753(1995).

[2]
RN SEQUENCE FROM N.A.
RP STRAIN=NM 9999;
RC Farmer A.D.;

RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U34166; AAA91713.1; -;
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1 494

FT NON_TER 494 494
SQ SEQUENCE 494 AA; 55063 MW; 735F7C2B86CAC58F CRC64;
Query Match 100.0%; Score 55; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AQIFNKPYW 9
Db 295 AQIFNKPYW 303
RESULT 18
Q81008 PRELIMINARY; PRT; 494 AA.
AC Q81008;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
LN L1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 7574;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
RT populations characterized by nucleotide sequence analysis of the B6,
RT L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 7574;
RA Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U34187; AAA91734.1; -;
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 494
SQ SEQUENCE 494 AA; 55008 MW; E55B8D4EF109F59E CRC64;
Query Match 100.0%; Score 55; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AQIFNKPYW 9
Db 295 AQIFNKPYW 303
RESULT 19
Q80999 PRELIMINARY; PRT; 494 AA.
AC Q80999;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
LN L1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.

OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NM T455;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
RT populations characterized by nucleotide sequence analysis of the B6,
RT L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NM T455;
RA Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U34169; AAA91716.1; -;
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 494
SQ SEQUENCE 494 AA; 55051 MW; 8D5F6C4CCDD5E5D8 CRC64;
Query Match 100.0%; Score 55; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AQIFNKPYW 9
Db 295 AQIFNKPYW 303
RESULT 20
Q81009 PRELIMINARY; PRT; 494 AA.
AC Q81009;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
LN L1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 7587;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
RT populations characterized by nucleotide sequence analysis of the B6,
RT L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 7587;
RA Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U34188; AAA91735.1; -;
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.

```

FT  NON_TER 1 1
FT  NON_TER 494 494
SQ  SEQUENCE 494 AA; 55094 MW; ECB31B09B8FBAE3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 494;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQTENKPYW 9
    |||||
    |||||
Db 295 AQTENKPYW 303

RESULT 21
ID Q81012 PRELIMINARY; PRT; 494 AA.
AC Q81012;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=OR 9237;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
RT populations characterized by nucleotide sequence analysis of the E6,
RT L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753 (1995).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=OR 9237;
RA Farmer A.D.;
RA Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U34193; AAA91740.1; -
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR Prodom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1 1
FT NON_TER 494 494
SQ SEQUENCE 494 AA; 55021 MW; B2D7AC39D176C3C5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 494;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQTENKPYW 9
    |||||
    |||||
Db 295 AQTENKPYW 303

RESULT 22
ID Q81001 PRELIMINARY; PRT; 494 AA.
AC Q81001;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

```

```

OC Papillomavirus.
OX NCBI_TaxID=10566;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=OR 1905;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
RT populations characterized by nucleotide sequence analysis of the E6,
RT L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753 (1995).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=OR 1905;
RA Farmer A.D.;
RA Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U34172; AAA91719.1; -
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR Prodom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1 1
FT NON_TER 494 494
SQ SEQUENCE 494 AA; 55094 MW; ECB31B09B8FBAE3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 494;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQTENKPYW 9
    |||||
    |||||
Db 295 AQTENKPYW 303

RESULT 23
ID Q90204 PRELIMINARY; PRT; 494 AA.
AC Q90204;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus type 16.
OS Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10561;
RN (1)
RP SEQUENCE FROM N.A.
RC MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
RT populations characterized by nucleotide sequence analysis of the E6,
RT L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753 (1995).
RN (2)
RP SEQUENCE FROM N.A.
RC Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U34173; AAA91720.1; -
DR EMBL; U34175; AAA91722.1; -
DR EMBL; U34178; AAA91725.1; -
DR EMBL; U34179; AAA91726.1; -
DR EMBL; U34168; AAA91715.1; -
DR EMBL; U34171; AAA91718.1; -
DR EMBL; U34180; AAA91727.1; -
DR EMBL; U34191; AAA91738.1; -
DR EMBL; U34192; AAA91739.1; -

```

DR EMBL; U34182; AAA91729.1; -
 DR EMBL; U34185; AAA91732.1; -
 DR GO; GO:0019028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR002210; PV capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR Prodom; PD000544; PV capsid_L1; 1.
 DR NON_TER 1
 FT NON_TER 494
 SO SEQUENCE 494 AA; 55021 MW; A737AC39CB6CC3DF CRC64;

Query Match 100.0%; Score 55; DB 12; Length 494;
 Best Local Similarity 100.0%; Pred. No. 0.052;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 A0IFNKPYW 9
 Db 295 A0IFNKPYW 303

RESULT 24
 ID 081000 PRELIMINARY; PRT; 494 AA.
 AC 081000;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Late major capsid protein L1 (Fragment).
 GN L1.
 OS Human papillomavirus.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OC NCBI_TaxID=10566;
 RX MEDLINE=96079021; PubMed=7494284;
 RX Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
 RA Jensen S.A.;
 RA "Human papillomavirus type 16 variant lineages in United States
 RT populations characterized by nucleotide sequence analysis of the B6,
 RT L2, and L1 coding segments.";
 RL J. Virol. 69:7743-7753(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NM T529;
 RA Farmer A.D.;
 RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U34170; AAA91717.1; -
 DR GO; GO:0019028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR002210; PV capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR Prodom; PD000544; PV capsid_L1; 1.
 DR NON_TER 1
 FT NON_TER 494
 SO SEQUENCE 494 AA; 55063 MW; 735F7C2B86CAC58F CRC64;

Query Match 100.0%; Score 55; DB 12; Length 494;
 Best Local Similarity 100.0%; Pred. No. 0.052;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 A0IFNKPYW 9
 Db 295 A0IFNKPYW 303

RESULT 25

ID 012343 PRELIMINARY; PRT; 495 AA.
 AC 012343;
 DT 01-JUL-1997 (TRENBLrel. 04, Created)
 DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Late major capsid protein L1 (Fragment).
 OS Human papillomavirus type 16.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OC NCBI_TaxID=10581;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PC-8;
 RA Tomesello M.L., Buongiorno F.M., Meglio A., Buongiorno L.,
 RA Beth-Girardo E., Girardo G.;
 RT "Sequence variations and viral genomic state of human papillomavirus
 RT type 16 in penile carcinomas from Ugandan patients.";
 RL J. Gen. Virol. 78:2199-2208(1997).
 DR EMBL; AF003031; AAB70748.1; -
 DR GO; GO:0019028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR002210; PV capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR Prodom; PD000544; PV capsid_L1; 1.
 DR NON_TER 1
 FT NON_TER 495
 SO SEQUENCE 495 AA; 55133 MW; 427DAFC2B42E1EDD CRC64;

Query Match 100.0%; Score 55; DB 12; Length 495;
 Best Local Similarity 100.0%; Pred. No. 0.052;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 A0IFNKPYW 9
 Db 298 A0IFNKPYW 306

RESULT 26
 ID 081017 PRELIMINARY; PRT; 504 AA.
 AC 081017;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE L1 capsid protein.
 OS Human papillomavirus type 31.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OC NCBI_TaxID=10585;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MS2;
 RX MEDLINE=96130211; PubMed=8553573;
 RX Icenogle J.P., Clancy K.A., Lin S.Y.;
 RA "Sequence variation in the capsid protein genes of human
 RT papillomavirus type 16 and type 31.";
 RL Virology 214:664-669(1995).
 DR EMBL; U37410; AAA92894.1; -
 DR GO; GO:0019028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR002210; PV capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR Prodom; PD000544; PV capsid_L1; 1.
 SO SEQUENCE 504 AA; 56335 MW; 0A3970E8631E2FA7 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 504;
 Best Local Similarity 100.0%; Pred. No. 0.053;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A0IFNKPYW 9
 |||||
 DB 305 A0IFNKPYW 312

RESULT 27

Q9WH57 PRELIMINARY; PRT; 505 AA.
 AC Q9WH57;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
 DE Capsid protein.
 GN LI.
 OS Human papillomavirus type 16.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10581;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GU2;
 RX MEDLINE=99252179; PubMed=10233949;
 RA White W.I., Wilson S.D., Palmer-Hill F.J., Woods R.M., Ghim S.J.,
 RA Hewitt L.A., Goldman D.M., Burke S.J., Jensen A.B., Koenig S.,
 RA Suzich J.A.;
 RT "Characterization of a major neutralizing epitope on human
 RT papillomavirus type 16 LI."
 RL J. Virol. 73:4882-4889(1999).
 DR EMBL; AF134178; AAD22113.1; -;
 DR GO; GO:0019028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR002210; PV_capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidLI.
 DR Prodom; PD000544; PV_capsid_L1; 1.
 SQ SEQUENCE 505 AA; 56245 MW; 37E59806BF1AA5F CRC64;

Query Match 100.0%; Score 55; DB 12; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.053;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A0IFNKPYW 9
 |||||
 DB 304 A0IFNKPYW 312

RESULT 28

Q9WL05 PRELIMINARY; PRT; 505 AA.
 AC Q9WL05;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
 DE Late major capsid protein.
 GN LI.
 OS Human papillomavirus type 16.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10581;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Phi1;
 RC Touze A., El-Mehdaoui S., Munoz N., Mouglin C., Couraget P.;
 RT "Papillomavirus type 16 full length sequence of LI gene."
 RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF043287; AAC09293.1; -;
 DR GO; GO:0019028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR002210; PV_capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidLI.
 DR Prodom; PD000544; PV_capsid_L1; 1.
 SQ SEQUENCE 505 AA; 56332 MW; C348AC9308D795ED CRC64;

Query Match 100.0%; Score 55; DB 12; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.053;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A0IFNKPYW 9
 |||||
 DB 304 A0IFNKPYW 312

RESULT 29

Q9WH56 PRELIMINARY; PRT; 505 AA.
 AC Q9WH56;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
 DE Capsid protein.
 GN LI.
 OS Human papillomavirus type 16.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10581;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GU1;
 RX MEDLINE=99252179; PubMed=10233949;
 RA White W.I., Wilson S.D., Palmer-Hill F.J., Woods R.M., Ghim S.J.,
 RA Hewitt L.A., Goldman D.M., Burke S.J., Jensen A.B., Koenig S.,
 RA Suzich J.A.;
 RT "Characterization of a major neutralizing epitope on human
 RT papillomavirus type 16 LI."
 RL J. Virol. 73:4882-4889(1999).
 DR EMBL; AF134177; AAD22112.1; -;
 DR GO; GO:0019028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR002210; PV_capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidLI.
 DR Prodom; PD000544; PV_capsid_L1; 1.
 SQ SEQUENCE 505 AA; 56260 MW; 435B61B6624BED85 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.053;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A0IFNKPYW 9
 |||||
 DB 304 A0IFNKPYW 312

RESULT 30

Q81015 PRELIMINARY; PRT; 505 AA.
 AC Q81015;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
 DE LI capsid protein.
 OS Human papillomavirus type 16.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10581;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1194;
 RX MEDLINE=96130211; PubMed=8553573;
 RX Icenogle J.P., Clancy K.A., Lin S.Y.;

RT "Sequence variation in the capsid protein genes of human
 RT Papillomavirus type 16 and type 31.";
 RL Virology 214:664-669(1995).
 DR EMBL; U37217; AAA92892.1; -
 DR GO; GO:0019028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR002210; PV_capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 DR PRODOM; PD000544; PV_capsid_L1; 1.
 SQ SEQUENCE 505 AA; 56344 MW; 52D2D06922836F68 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.053;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPW 9
 DB 304 AQIFNKPW 312

RESULT 31
 Q9WPH4 PRELIMINARY; PRT; 505 AA.
 AC Q9WPH4;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Capsid protein.
 OS Human papillomavirus type 16.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10581;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang L.L., Song G.X.;
 RT "The L1 Variant of Human Papillomavirus Type 16 in Condyloma
 RT Acuminatum Tissue from one Chinese Woman."
 RT Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF140365; AAD38523.1; -
 DR PDB; 1DZL; 25-AUG-00
 DR GO; GO:0019028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR002210; PV_capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 DR PRODOM; PD000544; PV_capsid_L1; 1.
 SQ SEQUENCE 505 AA; 56179 MW; 7526CECACE6B0E4 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.053;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPW 9
 DB 304 AQIFNKPW 312

RESULT 32
 Q92282 PRELIMINARY; PRT; 505 AA.
 ID Q92282;
 AC Q92282;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Late major capsid protein.
 GN L1.
 OS Human papillomavirus type 16.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10581;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang L.L., Song G.X., Si J.Y.;
 RT "Human papillomavirus type 16 late major capsid gene (L1)."
 RT Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF084952; AAC61736.1; -
 DR GO; GO:0019028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR002210; PV_capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 DR PRODOM; PD000544; PV_capsid_L1; 1.
 SQ SEQUENCE 505 AA; 56308 MW; 852BA3B3C09FD5B8 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.053;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPW 9
 DB 304 AQIFNKPW 312

RESULT 33
 Q9WH55 PRELIMINARY; PRT; 505 AA.
 AC Q9WH55;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Capsid protein.
 OS Human papillomavirus type 16.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10581;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Rocheester;
 RC MEDLINE=98105733; PubMed=9444988;
 RX White W.I., Wilson S.D., Bonnez W., Rose R.C., Koenig S., Suzich J.A.;
 RT "In vitro infection and type-restricted antibody-mediated
 RT neutralization of authentic human papillomavirus type 16.";
 RT J. Virol. 72:959-964 (1998).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rocheester;
 RX MEDLINE=98252179; PubMed=10233949;
 RA White W.I., Wilson S.D., Palmer-Hill F.J., Woods R.M., Chim S.J.,
 RA Hewitt L.A., Goldman D.M., Burke S.J., Jensen A.B., Koenig S.,
 RA Suzich J.A.;
 RT "Characterization of a major neutralizing epitope on human
 RT papillomavirus type 16 L1.";
 RT J. Virol. 73:4882-4889(1999).
 RL [3]
 DR EMBL; AF134175; AAD22111.1; -
 DR GO; GO:0019028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR002210; PV_capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 DR PRODOM; PD000544; PV_capsid_L1; 1.
 SQ SEQUENCE 505 AA; 56350 MW; E32A787B49B1ED5B CRC64;

Query Match 100.0%; Score 55; DB 12; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.053;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPW 9
 DB 304 AQIFNKPW 312

```

RESULT 34
ID Q9WLO6 PRELIMINARY; PRT; 505 AA.
AC Q9WLO6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Late major capsid protein (late major capsid protein L1).
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Gen32;
RA Touze A., El-Mehdaoui S., Munoz N., Mouglin C., Couraquet P.;
RT "Papillomavirus type 16 full length sequence of L1 gene.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=22645187; PubMed=12759944;
RA Guo H.X., Ling H., Zhang F.M., Zhong Z.H., Wang W.G., Guo S.Y., Lou G.,
RT "Cloning and sequencing of human papillomavirus 16 L1 gene from
RT cervical carcinoma tissues of Chinese women.";
RL Chung Hua Shih Yen Ho Lin Chuang Ping Tu Hsueh Tsa Chih
RL 13:17-19(1999).
RN (3)
RP SEQUENCE FROM N.A.
RA Gu H.X., Ling H., Zhang F.M., Zhong Z.H., Wang W.G., Guo S.Y.,
RA Zhuang M., Qi Z.Y., Lin D.H., Wei L.L., Shang Q.L., Ma P.L., Xiao P.,
RA Guo C.L., Li H.L., Huang J.L., Li X.R., Bao G.Y.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN (4)
RP SEQUENCE FROM N.A.
RA Varani A.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF043386; AAC09292.1; -
DR EMBL; AF393502; AAK7000.1; -
DR EMBL; AY177679; AAO19439.1; -
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR002210; FV_Capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPV_Capsid_L1.
DR PRODOM; PD000544; PV_capsid_L1; 1.
SQ SEQUENCE 505 AA; 56278 MW; D426CEC3DC6B0E2 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=85246220; PubMed=2990099;
RA Seedorf K., Krammer G., Durst M., Subai S., Rowekamp W.G.;
RT "Human papillomavirus type 16 DNA sequence.";
RL Virology 145:181-185(1985).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=87741358; PubMed=3029430;
RA Baker C.C., Phelps W.C., Lindgren V., Braun M.J., Gonda M.A.,
RA Howley P.M.;
RT "Structural and transcriptional analysis of human papillomavirus type
RT 16 sequences in cervical carcinoma cell lines.";
RL J. Virol. 61:962-971(1987).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE=87293897; PubMed=3039733;
RA el Awady M.K., Kaplan J.B., O'Brien S.J., Burk R.D.;
RT "Molecular analysis of integrated human papillomavirus 16 sequences in
RT the cervical cancer cell line SiHa.";
RL Virology 159:389-398(1987).
RN (4)
RP SEQUENCE FROM N.A.
RX MEDLINE=92194439; PubMed=1312620;
RA Chan S.Y., Ho L., Ong C.K., Chow V., Drescher B., Durst M.,
RA ter Meulen J., Villa L., Luande J., Mgaya H.N., Bernard H.-U.;
RT "Molecular variants of human papillomavirus type 16 from four
RT continents suggest ancient pandemic spread of the virus and its
RT coevolution with humankind.";
RL J. Virol. 66:2057-2066(1992).
RN (5)
RP SEQUENCE FROM N.A.
RA Meisner J.D.;
RT "Characterization of the HPV variants in Caski, SiHa, and HeLa
RT cervical carcinoma cell lines.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN (6)
RP SEQUENCE FROM N.A.
RC STRAIN=16M12E;
RA Flores E.R., Nelson J.H., Lambert P.F.;
RT "Establishment of the human papillomavirus life cycle in an
RT immortalized human foreskin keratinocyte cell line.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF001600; AAC31789.1; -
DR EMBL; AF125673; AAD33259.1; -
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR002210; FV_Capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPV_Capsid_L1.
DR PRODOM; PD000544; PV_capsid_L1; 1.
DR VARIANT 292 292 A -> T.
SQ SEQUENCE 531 AA; 59473 MW; 773491682106CC20 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 531;
Best Local Similarity 100.0%; Pred. No. 0.056;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RA Teraf M., Ma Z., Burk R.D.;
RT "Cloning and Sequencing of Non-European Human Papillomavirus (HPV)
RT Variants Complete Genomes from Patients with Cervical Cancer by an
RT Overlapping PCR Method.";
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF472508; AAO15704.1; -.
DR GO: GO:0019028; F:structural molecule activity; IEA.
DR InterPro: IPR002210; PV_capsid_L1.
DR Pfam: PF00500; late_protein_L1; 1.
DR PRINTS: PR00865; HPVcapsidL1.
DR PRODOM: PD000544; PV_capsid_L1; 1.
SQ SEQUENCE 531 AA; 59548 MW; 320626E6E4F6EFA3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 531;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
DB 330 AQIFNKPYW 338

RESULT 37
Q8BSN5 PRELIMINARY; PRT; 531 AA.
AC Q8BSN5;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Putative major capsid protein L1.
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RA Teraf M., Ma Z., Burk R.D.;
RT "Cloning and Sequencing of Non-European Human Papillomavirus (HPV)
RT Variants Complete Genomes from Patients with Cervical Cancer by an
RT Overlapping PCR Method.";
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF472509; AAO15712.1; -.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002210; PV_capsid_L1.
DR Pfam: PF00500; late_protein_L1; 1.
DR PRINTS: PR00865; HPVcapsidL1.
DR PRODOM: PD000544; PV_capsid_L1; 1.
SQ SEQUENCE 531 AA; 59541 MW; F1C433C8ADA313AA CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 531;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
DB 330 AQIFNKPYW 338

RESULT 38
Q8OML8 PRELIMINARY; PRT; 531 AA.
AC Q8OML8;

```

```

DT 01-JUN-2003 (TREMblrel. 24, Created)
DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Asian-American variant;
RA Teraf M., Burk R.D.;
RT "Human papillomavirus type 16 Asian-American variant.";
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF402678; AAO85415.1; -.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002210; PV_capsid_L1.
DR Pfam: PF00500; late_protein_L1; 1.
DR PRINTS: PR00865; HPVcapsidL1.
DR PRODOM: PD000544; PV_capsid_L1; 1.
SQ SEQUENCE 531 AA; 59517 MW; 72E7F90A83D48186 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 531;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
DB 330 AQIFNKPYW 338

RESULT 39
O56951 PRELIMINARY; PRT; 565 AA.
ID O56951;
AC O56951;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE L1 protein.
GN L1.
OS Human papillomavirus type 77.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=69986;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98118461; PubMed=9454709;
RA Delius H., Seegling B., Bergmann K., Shamanin V., de Villiers E.M.;
RT "The genomes of three of four novel HPV types, defined by differences
RT of their L1 genes, show high conservation of the E7 gene and the
RT URR.";
RL Virology 240:359-365 (1998).
DR EMBL: Y15175; CAAT5468.1; -.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002210; PV_capsid_L1.
DR Pfam: PF00500; late_protein_L1; 1.
DR PRINTS: PR00865; HPVcapsidL1.
DR PRODOM: PD000544; PV_capsid_L1; 1.
SQ SEQUENCE 565 AA; 63377 MW; 393B5933F9480E67 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 565;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
DB 366 AQIFNKPYW 374

```

RESULT 40

Q84304 PRELIMINARY; PRT; 44 AA.
 ID Q84304
 AC Q84304;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE L1 (Fragment).
 GN L1.
 OS Human papillomavirus type 10.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_Taxid=10603;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92407963; PubMed=1326639;
 RA Chan S.-Y., Bernard H.-U., Ong C.-K., Chan S.-P., Birgit H.,
 RA Delius H.;
 RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and
 RT variants: A showcase for the molecular evolution of DNA viruses.";
 RL J. Virol. 66:5714-5725(1992).
 DR EMBL; M96280; AAA47019.1; -.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR002210; PV_capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 FT NON_TER 1
 SQ SEQUENCE 44 AA; 5088 MW; 0A5860E6447B1B4B CRC64;

Query Match 96.4%; Score 53; DB 12; Length 44;
 Best Local Similarity 88.9%; Pred. No. 0.01; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AOIFNKPYW 9
 ||:|||||
 Db 3 AOIFNKPYW 11

Search completed: March 3, 2004, 10:24:46
 Job time : 38.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 10:22:54 ; Search time 16.5 Seconds

(Without alignments)
28.160 Million cell updates/sec

Title: US-09-980-064-2

Perfect score: 49

Sequence: 1 AGVDNRECT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiletest.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	15	4	US-09-947-372A-17
2	49	100.0	26	2	US-08-943-583-3
3	49	100.0	27	2	US-08-943-583-2
4	49	100.0	497	4	US-09-413-611A-6
5	49	100.0	499	4	US-09-413-611A-8
6	49	100.0	505	3	US-08-944-368A-2
7	49	100.0	505	4	US-09-820-764-2
8	49	100.0	505	4	US-09-986-118A-2
9	49	100.0	505	4	US-09-824-017-2
10	49	100.0	531	4	US-09-520-822A-1
11	39	79.6	20	1	US-07-678-974D-12
12	39	79.6	20	1	US-08-945-168-17
13	36	73.5	20	1	US-07-678-974D-11
14	36	73.5	20	2	US-08-945-168-16
15	36	73.5	26	2	US-08-943-583-4
16	35	71.4	126	1	US-08-451-947-57
17	35	71.4	126	2	US-08-424-826A-57
18	35	71.4	126	3	US-08-928-694-57
19	35	71.4	126	4	US-08-450-842-57
20	35	71.4	126	4	US-08-451-390-57
21	35	71.4	126	5	PCT-US91-06950-57
22	34	63.4	134	2	US-08-847-724-1
23	34	63.4	169	4	US-09-134-001C-4000
24	34	63.4	647	4	US-09-540-236-2099
25	34	63.4	1015	1	US-08-537-210A-1
26	34	63.4	1015	3	US-09-113-825-1
27	34	63.4	1078	1	US-08-264-534-32

28	34	69.4	1078	1	US-08-083-590A-11	Sequence 11, Appl
29	34	69.4	1078	1	US-08-465-500-32	Sequence 32, Appl
30	34	69.4	1078	2	US-08-346-128-32	Sequence 32, Appl
31	34	69.4	1078	3	US-08-532-384-11	Sequence 11, Appl
32	34	69.4	1078	3	US-08-893-828-32	Sequence 32, Appl
33	34	69.4	1078	3	US-08-185-432-16	Sequence 16, Appl
34	34	69.4	2471	1	US-08-083-590A-19	Sequence 19, Appl
35	34	69.4	2471	3	US-08-532-384-19	Sequence 19, Appl
36	34	69.4	2471	4	US-08-899-232-1	Sequence 1, Appl
37	34	69.4	2556	1	US-08-083-590A-20	Sequence 20, Appl
38	34	69.4	2556	3	US-08-532-384-20	Sequence 20, Appl
39	33	67.3	157	4	US-09-540-236-2507	Sequence 2507, Ap
40	33	67.3	984	2	US-08-673-789-9	Sequence 9, Appl
41	33	67.3	984	2	US-08-449-645A-19	Sequence 19, Appl
42	33	67.3	984	2	US-08-702-367A-19	Sequence 19, Appl
43	33	67.3	984	5	PCT-US95-04681-19	Sequence 19, Appl
44	32	65.3	26	2	US-08-943-583-5	Sequence 5, Appl
45	32	65.3	26	2	US-08-943-583-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-947-372A-17
Sequence 17, Application US/09947372A
Patent No. 6613557
GENERAL INFORMATION:
APPLICANT: FRAZER, IAN
TITLE OF INVENTION: PAPILLOMAVIRUS VACCINE
FILE REFERENCE: 065064/0137
CURRENT APPLICATION NUMBER: US/09/947,372A
CURRENT FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 08/185,928
PRIOR FILING DATE: 1994-01-19
PRIOR APPLICATION NUMBER: PCT/AU92/02184
PRIOR FILING DATE: 1992-07-20
PRIOR APPLICATION NUMBER: AU PK7322
PRIOR FILING DATE: 1991-07-19
NUMBER OF SEQ ID NOS: 66
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 17
LENGTH: 15
TYPE: PRT
ORGANISM: Human papillomavirus type 16
US-09-947-372A-17
Query Match 100.0%; Score 49; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGVDNRECT 9
DB 5 AGVDNRECT 13
RESULT 2
US-08-943-583-3
Sequence 3, Application US/08943583
Patent No. 5922588
GENERAL INFORMATION:
APPLICANT: Ludmerer, Steven
TITLE OF INVENTION: HPV 16 VIRUS-LIKE PARTICL
TITLE OF INVENTION: ES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,583
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/026,763
; FILING DATE: 04-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gieseler, Joanne M
; REGISTRATION NUMBER: 32,838
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-3046
; TELEFAX: 732-594-4720
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
;
US-08-943-583-3

Query Match          100.0%; Score 49; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0066;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGVDNRECI 9
Db      17 AGVDNRECI 25

RESULT 3
US-08-943-583-2
; Sequence 2, Application US/08943583
; Patent No. 5922588
; GENERAL INFORMATION:
; APPLICANT: Ludmerer, Steven
; TITLE OF INVENTION: HPV 16 VIRUS-LIKE PARTICL
; TITLE OF INVENTION: ES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,583
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/026,763
; FILING DATE: 04-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gieseler, Joanne M
; REGISTRATION NUMBER: 32,838
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-3046
; TELEFAX: 732-594-3046
;

```

```

;
; TELEFAX: 732-594-4720
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
;
US-08-943-583-2

Query Match          100.0%; Score 49; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.0068;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGVDNRECI 9
Db      16 AGVDNRECI 24

RESULT 4
US-09-413-611A-6
; Sequence 6, Application US/09413611A
; Patent No. 6380364
; GENERAL INFORMATION:
; APPLICANT: Mueller, Martin
; APPLICANT: Kast, Wiljbe
; APPLICANT: Nieland, John
; APPLICANT: Veldeers, Markwin
; TITLE OF INVENTION: Chimeric Biotin-Binding Papillomavirus Protein
; FILE REFERENCE: 202325
; CURRENT APPLICATION NUMBER: US/09/413,611A
; CURRENT FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 60/109,510
; PRIOR FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:biotin-binding
;
US-09-413-611A-6

Query Match          100.0%; Score 49; DB 4; Length 497;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGVDNRECI 9
Db      139 AGVDNRECI 147

RESULT 5
US-09-413-611A-8
; Sequence 8, Application US/09413611A
; Patent No. 6380364
; GENERAL INFORMATION:
; APPLICANT: Mueller, Martin
; APPLICANT: Kast, Wiljbe
; APPLICANT: Nieland, John
; APPLICANT: Veldeers, Markwin
; TITLE OF INVENTION: Chimeric Biotin-Binding Papillomavirus Protein
; FILE REFERENCE: 202325
; CURRENT APPLICATION NUMBER: US/09/413,611A
; CURRENT FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 60/109,510
; PRIOR FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
;

```

SEQ ID NO 8
LENGTH: 499
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: biotin-binding
OTHER INFORMATION: papillomavirus protein 169
US-09-413-611A-8

Query Match 100.0%; Score 49; DB 4; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
|||||
Db 139 AGVDNRECI 147

RESULT 6
US-08-944-368A-2
Sequence 2, Application US/08944368A
Patent No. 6228368
GENERAL INFORMATION:

APPLICANT: Gissman, et al.
TITLE OF INVENTION: Papilloma Virus Capsomere Vaccine
TITLE OF INVENTION: Formulations and Methods of Use
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,368A
FILING DATE:

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27013/34028
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-944-368A-2

Query Match 100.0%; Score 49; DB 3; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
|||||
Db 139 AGVDNRECI 147

RESULT 7
US-09-820-764-2
Sequence 2, Application US/09820764
Patent No. 6352696
GENERAL INFORMATION:

APPLICANT: BURGER, Alexander
HALLER, Michael
TITLE OF INVENTION: PAPILLOMA VIRUS CAPSOMERE VACCINE
FORMULATIONS AND METHODS OF USE
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/820,764
FILING DATE: 30-Mar-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/026,896
FILING DATE: 20-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Sandercock, Colin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 37067/102
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-820-764-2

Query Match 100.0%; Score 49; DB 4; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
|||||
Db 139 AGVDNRECI 147

RESULT 8
US-09-986-118A-2
Sequence 2, Application US/09986118A
Patent No. 6562351
GENERAL INFORMATION:

APPLICANT: BURGER, Alexander
HALLER, Michael
TITLE OF INVENTION: PAPILLOMA VIRUS CAPSOMERE VACCINE
FORMULATIONS AND METHODS OF USE
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/986,118A

```

;
; FILING DATE: 07-No. 6562351-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/026,896
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercock, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 37067/102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 505 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-986-118A-2

Query Match      100.0%; Score 49; DB 4; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECT 9
DB 139 AGVDNRECT 147

RESULT 9
US-09-824-017-2
; Sequence 2, Application US/09824017
; Patent No. 6649167
; GENERAL INFORMATION:
; APPLICANT: BURGER, Alexander
; TITLE OF INVENTION: PAPILLOMA VIRUS CAPSOMERE VACCINE
; FORMULATIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/824,017
; FILING DATE: 03-Apr-2001
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/026,896
; FILING DATE: 1998-02-20
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercock, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 37067/102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 505 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-824-017-2
; Sequence 1, Application US/09520822A
; Patent No. 6551597
; GENERAL INFORMATION:
; APPLICANT: Harrison, Stephen
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
; FILE REFERENCE: Harvard/Harrison 12687/1120
; CURRENT APPLICATION NUMBER: US/09/520,822A
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/125208
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/148544
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-09-520-822A-1

Query Match      100.0%; Score 49; DB 4; Length 531;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECT 9
DB 165 AGVDNRECT 173

RESULT 11
US-07-678-974D-12
; Sequence 12, Application US/07678974D
; Patent No. 5629146
; GENERAL INFORMATION:
; APPLICANT: DILLNER, JOAKIM
; TITLE OF INVENTION: METHOD FOR DETECTION OF HUMAN PAPILLOMAVIRUS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERMAN & AISENBERG
; STREET: 1730 RHODE ISLAND AVENUE, N.W.,
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-3186
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/678,974D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: AISENBERG, Irwin M.
; REGISTRATION NUMBER: 19,007
```

REFERENCE/DOCKET NUMBER: SG19171
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-293-1404
TELEFAX: 202-872-0493
TELEX: 440 069 AIS UI
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-07-678-974D-12

Query Match 79.6%; Score 39; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 VDNRECI 9
Db 1 VDNRECI 7

RESULT 12
US-08-945-168-17
Sequence 17, Application US/08945168
Patent No. 5989548
GENERAL INFORMATION:
APPLICANT: DILLNER, JOAKIM
TITLE OF INVENTION: PEPTIDE-BASED VACCINE AGAINST PAPILLOMA
TITLE OF INVENTION: VIRUS
NUMBER OF SEQUENCES: 117
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P. C.
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945.168
FILING DATE: 18-DEC-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SE96/00533
FILING DATE: 23-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9501512-9
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7752-0002-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-945-168-17

Query Match 79.6%; Score 39; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 VDNRECI 9
Db 1 VDNRECI 7

RESULT 13
US-07-678-974D-11
Sequence 11, Application US/07678974D
Patent No. 5629146
GENERAL INFORMATION:
APPLICANT: DILLNER, JOAKIM
TITLE OF INVENTION: METHOD FOR DETECTION OF HUMAN PAPILLOMAVIRUS
TITLE OF INVENTION:
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERMAN & AISENBERG
STREET: 1730 RHODE ISLAND AVENUE, N.W.,
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20036-3186
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/678.974D
FILING DATE: 25-JUN-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: AISENBERG, Irwin M.
REGISTRATION NUMBER: 19,007
REFERENCE/DOCKET NUMBER: SG19171
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-293-1404
TELEFAX: 202-872-0493
TELEX: 440 069 AIS UI
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-07-678-974D-11

Query Match 73.5%; Score 36; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGVDNRE 7
Db 14 AGVDNRE 20

RESULT 14
US-08-945-168-16
Sequence 16, Application US/08945168
Patent No. 5989548
GENERAL INFORMATION:
APPLICANT: DILLNER, JOAKIM
TITLE OF INVENTION: PEPTIDE-BASED VACCINE AGAINST PAPILLOMA
TITLE OF INVENTION: VIRUS
NUMBER OF SEQUENCES: 117
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P. C.
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.10
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,168
FILING DATE: 18-DEC-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SE96/00533
FILING DATE: 23-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9501512-9
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7752-0002-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-945-168-16

Query Match 73.5%; Score 36; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRE 7
|||||
Db 14 AGVDNRE 20

RESULT 15
US-08-943-583-4
Sequence 4, Application US/08943583
Patent No. 5922588
GENERAL INFORMATION:
APPLICANT: Ludmerer, Steven
TITLE OF INVENTION: HPV 16 VIRUS-LIKE PARTICL
TITLE OF INVENTION: ES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,583
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/026,763
FILING DATE: 04-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gieseler, Joanne M
REGISTRATION NUMBER: 32,838
REFERENCE/DOCKET NUMBER: 19803
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-3046

TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-943-583-4

Query Match 73.5%; Score 36; DB 2; Length 26;
Best Local Similarity 77.8%; Pred. No. 2;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
|||||
Db 17 AGVDNRECI 25

RESULT 16
US-08-451-947-57
Sequence 57, Application US/08451947
Patent No. 5702906
GENERAL INFORMATION:
APPLICANT: GENENTECH, INC.
APPLICANT: ROSENTHAL, ARNON
TITLE OF INVENTION: NOVEL NEUROTROPIC FACTOR
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,947
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/426419
FILING DATE: 19-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/030013
FILING DATE: 22-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648482
FILING DATE: 31-JAN
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/587707
FILING DATE: 1991
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: 666P2C1D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-451-947-57

Query Match 71.4%; Score 35; DB 1; Length 126;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVDNREC 8
||| |||
Db 80 GVDNREC 86

RESULT 17
US-08-424-826A-57
; Sequence 57, Application US/08424826A
; Patent No. 5830658
; GENERAL INFORMATION:
; APPLICANT: Rosenthal, Arnon
; TITLE OF INVENTION: NOVEL NEUTROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,826A
; FILING DATE: 19-Apr-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/240387
; FILING DATE: 10-May-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/648482
; FILING DATE: 31-Jan-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/587707
; FILING DATE: 25-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P0666P1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8674
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-424-826A-57

Query Match 71.4%; Score 35; DB 2; Length 126;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVDNREC 8
||| |||
Db 80 GVDNREC 86

RESULT 18
US-08-928-694-57
; Sequence 57, Application US/08928694
; Patent No. 6037320
; GENERAL INFORMATION:
; APPLICANT: ROSENTHAL, ARNON
; TITLE OF INVENTION: NOVEL NEUTROTROPHIC FACTOR

NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,694
; FILING DATE: 12-Sep-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/451947
; FILING DATE: 26-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/426419
; FILING DATE: 19-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/030013
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/648482
; FILING DATE: 31-JAN
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/587707
; FILING DATE: 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P0666P2C1D2C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-928-694-57

Query Match 71.4%; Score 35; DB 3; Length 126;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVDNREC 8
||| |||
Db 80 GVDNREC 86

RESULT 19
US-08-450-842-57
; Sequence 57, Application US/08450842
; Patent No. 6506728
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; TITLE OF INVENTION: NOVEL NEUTROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,842
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA: 08/426419
APPLICATION NUMBER: 08/426419
FILING DATE: 19-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/030013
FILING DATE: 22-MAR-1993
PRIOR APPLICATION DATA: 07/648482
APPLICATION NUMBER: 07/648482
FILING DATE: 31-JAN
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/587707
FILING DATE: 1991
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: 666P2C1D3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-450-842-57
```

```
Query Match 71.4%; Score 35; DB 4; Length 126;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2 GVDNREC 8
DB 80 GVDNREC 86
```

```
RESULT 20
US-08-451-390-57
Sequence 57, Application US/08451390
Patent No. 6566091
GENERAL INFORMATION:
APPLICANT: GENENTECH, INC.
TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSER: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,390
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/426419
FILING DATE: 19-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/030013
```

```
FILING DATE: 22-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648482
FILING DATE: 31-JAN
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/587707
FILING DATE: 1991
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: 666P2C1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-451-390-57
```

```
Query Match 71.4%; Score 35; DB 4; Length 126;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2 GVDNREC 8
DB 80 GVDNREC 86
```

```
RESULT 21
PCT-US91-06950-57
Sequence 57, Application PC/TUS9106950
GENERAL INFORMATION:
APPLICANT: GENENTECH, INC.
APPLICANT: ROSENTHAL, ARNON
TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSER: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/06950
FILING DATE: 19910924
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648482
APPLICATION NUMBER: 07/587707
ATTORNEY/AGENT INFORMATION:
NAME: Hensley, Max D.
REGISTRATION NUMBER: 27,043
REFERENCE/DOCKET NUMBER: 666P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US91-06950-57
```

Query Match 71.4%; Score 35; DB 5; Length 126;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVDNREC 8
Db 80 GVDNREC 86

RESULT 22
US-08-847-724-1
; Sequence 1, Application US/08847724
; Patent No. 595305
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Petichory, Joanne R.
; TITLE OF INVENTION: NOVEL HUMAN RETINOID BINDING PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,724
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0280 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 134 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: THYRN0202
; CLONE: 879706
; US-08-847-724-1

Query Match 69.4%; Score 34; DB 2; Length 134;
Best Local Similarity 71.4%; Pred. No. 27;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVDNREC 8
Db 77 GVDNREC 83

RESULT 23
US-09-134-001C-4000
; Sequence 4000, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4000
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4000

Query Match 69.4%; Score 34; DB 4; Length 169;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGVDNREC 9
Db 137 AGVDNREC 145

RESULT 24
US-09-540-236-2099
; Sequence 2099, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2099
; LENGTH: 647
; TYPE: PRT
; ORGANISM: M.catarrhalis
; US-09-540-236-2099

Query Match 69.4%; Score 34; DB 4; Length 647;
Best Local Similarity 62.5%; Pred. No. 1,4e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGVDNREC 8
Db 181 AGVDNREC 188

RESULT 25
US-08-537-210A-1
; Sequence 1, Application US/08537210A
; Patent No. 5780300
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Fortini, Mark
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: MANIPULATION OF NON-TERMINALLY
; TITLE OF INVENTION: DIFFERENTIATED CELLS USING THE NOTCH PATHWAY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/537,210A
FILING DATE: 29-SEP-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-027
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNITE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1015 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: hum N (Human No. 5780300ch 2)
LOCATION: 1155...2169
OTHER INFORMATION: Highly conserved ankyrin repeat
OTHER INFORMATION: region of No. 5780300ch
US-08-537-210A-1

Query Match 69.4%; Score 34; DB 1; Length 1015;
Best Local Similarity 57.1%; Pred. No. 2.2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VDNRECT 9
Db 473 IDNRQCV 479

RESULT 26
US-09-113-825-1
Sequence 1, Application US/09113825
GENERAL INFORMATION:
PATENT: 6149902
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Fortini, Mark
APPLICANT: Matsuno, Kenji
TITLE OF INVENTION: MANIPULATION OF NON-TERMINALLY
TITLE OF INVENTION: DIFFERENTIATED CELLS USING THE NOTCH PATHWAY
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/113,825
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/537,210
FILING DATE: 29-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-027
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864
TELEX: 66141 PENNITE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1015 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: hum N (Human No. 6149902ch 2)
LOCATION: 1155...2169
OTHER INFORMATION: Highly conserved ankyrin repeat
OTHER INFORMATION: region of No. 6149902ch
US-09-113-825-1

Query Match 69.4%; Score 34; DB 3; Length 1015;
Best Local Similarity 57.1%; Pred. No. 2.2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VDNRECT 9
Db 473 IDNRQCV 479

RESULT 27
US-08-264-534-32
Sequence 32, Application US/08264534
PATENT: 5648464
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon et al.
TITLE OF INVENTION: Human No. 5648464ch And Delta, Binding Domains
TITLE OF INVENTION: In Topolythmic Proteins, And Methods Based Thereon
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/264,534
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/695,189
FILING DATE: 03-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNITE
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-264-534-32

Query Match 69.4%; Score 34; DB 1; Length 1078;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 VDNRECI 9
Db 227 IDNRQCV 233

RESULT 28

US-08-083-590A-11
Sequence 11, Application US/08083590A
Patent No. 5786158
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
TITLE OF INVENTION: Nucleic Acids
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,590A
FILING DATE: 25-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-083-590A-11

Query Match 69.4%; Score 34; DB 1; Length 1078;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 VDNRECI 9
Db 227 IDNRQCV 233

RESULT 29

US-08-465-500-32
Sequence 32, Application US/08465500
Patent No. 5789195
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Muskavitch, Marc A.T.
APPLICANT: Fehon, Richard G.
APPLICANT: Rebay, Ilaria
APPLICANT: Blumheller, Cristine M.
APPLICANT: Shepard, Scott B.
TITLE OF INVENTION: HUMAN NOTCH AND DELTA, BINDING DOMAINS
TITLE OF INVENTION: IN TOPORYTHMIC PROTEINS, AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas

CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,500
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-465-500-32

Query Match 69.4%; Score 34; DB 1; Length 1078;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 VDNRECI 9
Db 227 IDNRQCV 233

RESULT 30

US-08-346-128-32
Sequence 32, Application US/08346128
Patent No. 5856441
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon et al.
TITLE OF INVENTION: Human No. 5856441ch And Delta, Binding Domains
TITLE OF INVENTION: In Toporythmic Proteins, And Methods Based Thereon
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/346,128
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/879,038
FILING DATE: 30-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-346-128-32

Query Match 69.4%; Score 34; DB 2; Length 1078;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNRECI 9
:||||:
DB 227 IDNRQCV 233

RESULT 31
US-08-532-384-11
Sequence 11, Application US/08532384
Patent No. 6083904
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
TITLE OF INVENTION: Nucleic Acids
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,384
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/083,590
FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-532-384-11

Query Match 69.4%; Score 34; DB 3; Length 1078;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNRECI 9
:||||:
DB 227 IDNRQCV 233

RESULT 32

US-08-893-828-32
Sequence 32, Application US/08893828
Patent No. 6090922
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Muskavitch, Marc A.T.
APPLICANT: Fehon, Richard G.
APPLICANT: Rebay, Ilaria
APPLICANT: Blumheller, Cristine M.
APPLICANT: Shepard, Scott B.
TITLE OF INVENTION: HUMAN NOTCH AND DELTA, BINDING DOMAINS
TITLE OF INVENTION: IN TOPOGRAPHIC PROTEINS, AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893,828
FILING DATE: 11-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-893-828-32

Query Match 69.4%; Score 34; DB 3; Length 1078;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNRECI 9
:||||:
DB 227 IDNRQCV 233

RESULT 33
US-08-185-432-16
Sequence 16, Application US/08185432
Patent No. 5750652
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Bussseau, Isabelle
APPLICANT: Diederich, Robert J.
APPLICANT: Xu, Tian
APPLICANT: Matsuno, Kenji
TITLE OF INVENTION: DETEX PROTEINS, NUCLEIC ACIDS, AND
TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.

ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,432
FILING DATE: 21-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2471 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-185-432-16

Query Match 69.4%; Score 34; DB 1; Length 2471;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNECI 9
DB 1627 IDNRQCV 1633

RESULT 34
US-08-083-590A-19
Sequence 19, Application US/08083590A
Patent No. 5786158
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
TITLE OF INVENTION: Nucleic Acids
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,590A
FILING DATE: 25-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 2471 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-083-590A-19

Query Match 69.4%; Score 34; DB 1; Length 2471;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNECI 9
DB 1627 IDNRQCV 1633

RESULT 35
US-08-532-384-19
Sequence 19, Application US/08532384
Patent No. 6083904
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
TITLE OF INVENTION: Nucleic Acids
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,384
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/083,590
FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 2471 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-532-384-19

Query Match 69.4%; Score 34; DB 3; Length 2471;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNECI 9
DB 1627 IDNRQCV 1633

RESULT 36
US-08-899-232-1
Sequence 1, Application US/08899232
Patent No. 6436650

GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Qi, Huilin
TITLE OF INVENTION: ACTIVATED FORMS OF NOTCH AND METHODS BASED THEREON
FILE REFERENCE: 7326-046
CURRENT APPLICATION NUMBER: US/08/899,232
CURRENT FILING DATE: 1997-07-23
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 1
LENGTH: 2471
TYPE: PRT
ORGANISM: Homo sapiens
US-08-899-232-1

Query Match 69.4%; Score 34; DB 4; Length 2471;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VNRECT 9
:||||:
DB 1627 IDNRQCV 1633

RESULT 37
US-08-893-590A-20
Sequence 20, Application US/08083590A
Patent No. 5786158
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
TITLE OF INVENTION: Nucleic Acids
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,590A
FILING DATE: 25-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 2556 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-893-590A-20

Query Match 69.4%; Score 34; DB 1; Length 2556;
Best Local Similarity 57.1%; Pred. No. 5.9e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 3 VNRECT 9
:||||:
DB 1627 IDNRQCV 1633

DB 1681 IDNRQCV 1687

RESULT 38
US-08-532-384-20
Sequence 20, Application US/08532384
Patent No. 6083904
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
TITLE OF INVENTION: Nucleic Acids
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,384
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/083,590
FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 2556 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-532-384-20

Query Match 69.4%; Score 34; DB 3; Length 2556;
Best Local Similarity 57.1%; Pred. No. 5.9e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VNRECT 9
:||||:
DB 1681 IDNRQCV 1687

RESULT 39
US-09-540-236-2507
Sequence 2507, Application US/09540236
Patent No. 6673910
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
FILE REFERENCE: 2709.2005-001
CURRENT APPLICATION NUMBER: US/09/540,236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 2507
LENGTH: 157
TYPE: PRT

; ORGANISM: M.catarrhalis
US-09-540-236-2507

Query Match 67.3%; Score 33; DB 4; Length 157;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 AGVDNRECT 9
||:|
Db 33 AGIDLRACI 41

RESULT 40

US-08-673-789-9
; Sequence 9, Application US/08673789
; Patent No. 5814479
; GENERAL INFORMATION:
; APPLICANT: ZHOU, RENPING; SCHULZ, NICHOLAS,
; APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOIDE,
; APPLICANT: GEORGE, F.
; TITLE OF INVENTION: BSK RECEPTOR LIKE
; TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
; TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORGAN & PINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/673,789
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/177,812
; FILING DATE: 04-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CAROL M. GRUPPI
; REGISTRATION NUMBER: 37,341
; REFERENCE/DOCKET NUMBER: 2026-4105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEEX: 421792
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 984
; TYPE: AMINO ACID
; STRANDEDNESS: UNKNOWN
; TOPOLOGY: UNKNOWN
; US-08-673-789-9

Query Match 67.3%; Score 33; DB 2; Length 984;
Best Local Similarity 55.6%; Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 AGVDNRECT 9
||:|
Db 945 AGIDTMECV 953

Search completed: March 3, 2004, 10:26:10
Job time : 17.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 09:46:53 ; Search time 50.5 Seconds
(without alignments)
50.355 Million cell updates/sec

Title: US-09-980-064-1
Perfect score: 55
Sequence: 1 AQLFNKPYW 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20048:*
8: geneseqp20058:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	9	AA46246	Abp47256 HPV type
2	55	100.0	20	ABP47256	Abp47256 HPV16 L1
3	55	100.0	497	AA71464	Aay71464 Chimeric
4	55	100.0	499	AA71465	Aay71465 Chimeric
5	55	100.0	504	AA98436	AA98436 Human pap
6	55	100.0	505	AA98436	AA98436 HPV1 L1
7	55	100.0	505	AA98436	AA98436 HPV1 L1
8	55	100.0	505	AA98436	AA98436 HPV1 L1
9	55	100.0	505	AA98436	AA98436 HPV1 L1
10	55	100.0	505	AA98436	AA98436 HPV1 L1
11	55	100.0	505	AA98436	AA98436 HPV1 L1
12	55	100.0	505	AA98436	AA98436 HPV1 L1
13	55	100.0	505	AA98436	AA98436 HPV1 L1
14	55	100.0	505	AA98436	AA98436 HPV1 L1
15	55	100.0	505	AA98436	AA98436 HPV1 L1
16	55	100.0	505	AA98436	AA98436 HPV1 L1
17	55	100.0	505	AA98436	AA98436 HPV1 L1
18	55	100.0	505	AA98436	AA98436 HPV1 L1
19	55	100.0	505	AA98436	AA98436 HPV1 L1
20	55	100.0	505	AA98436	AA98436 HPV1 L1
21	55	100.0	505	AA98436	AA98436 HPV1 L1
22	55	100.0	505	AA98436	AA98436 HPV1 L1
23	55	100.0	505	AA98436	AA98436 HPV1 L1
24	55	100.0	505	AA98436	AA98436 HPV1 L1
25	55	100.0	505	AA98436	AA98436 HPV1 L1

26	55	100.0	505	6	ADA27363	Ada27363 HPV-16 L1
27	55	100.0	505	7	ADA92541	Ada92541 HPV-16 L1
28	55	100.0	505	7	ADA14290	Ada14290 HPV-16 L1
29	55	100.0	505	7	AAE38615	AAE38615 HPV-16 L1
30	55	100.0	505	7	ADC26174	ADC26174 HPV16 L1
31	55	100.0	505	7	ADC26175	ADC26175 HPV16 L1
32	55	100.0	531	3	AA98436	AA98436 Human pap
33	55	100.0	531	3	AA98436	AA98436 Human pap
34	55	100.0	533	5	ABP77482	ABP77482 HPV16-L1/
35	55	100.0	533	5	ABP77481	ABP77481 HPV16-L1/
36	55	100.0	542	5	AAW47281	AAW47281 Papilloma
37	53	96.4	128	2	AAW47280	AAW47280 Papilloma
38	53	96.4	467	6	AAO16091	AAO16091 Human pap
39	53	96.4	467	6	AAO16092	AAO16092 Humanised
40	53	96.4	483	6	AAO16097	AAO16097 HPV L1/Ub
41	53	96.4	500	2	AAW10647	AAW10647 HPV11 mut
42	53	96.4	500	2	AAW24582	AAW24582 HPV6 L1 P
43	53	96.4	500	2	AAW15113	AAW15113 Human HPV
44	53	96.4	500	2	AAW15112	AAW15112 Human HPV
45	53	96.4	500	2	AAW15112	AAW15112 Human HPV

ALIGNMENTS

RESULT 1
ID AAB46246 standard; peptide; 9 AA.
AC AAB46246;
DT 04-APR-2001 (first entry)
DE HPV type 16 cytotoxic T-cell epitope SEQ ID NO 1.
KW Cytotoxic T cell; epitope; L1 protein; antiviral; antitumor; antigen; vaccine; tumor; protective immune response.
OS Human papillomavirus.
PN DE19925235-A1.
PD 07-DEC-2000.
PF 01-JUN-1999; 99DE-01025235.
PR 01-JUN-1999; 99DE-01025235.
PA (MED1-) MEDIGENE AG.
PI (DKFZ-) DKFZ DEUT KRBBRSCHUNGSGEZENTRUM.
PI Schaefer K, Faath S, Jochims I, Nieland J, Ossen W, WPI, 2001-062580/08.
PT New T cell epitopes from the papilloma virus L1 protein, useful for detecting or inducing an immune response, e.g. in vaccines.
PS Claim 1, Page 11, 26pp; German.
XX This invention describes novel T cell epitopes (A) of the human papilloma virus L1 protein. The invention also describes (a) compounds (I) containing an (A) provided it is not a natural L1 protein nor an exclusively N- or C-terminal deletion variant of natural L1; (b) nucleic acid (II) that encodes (A) or (I); (c) vector, particularly an expression vector, that contains (II); (d) cells that contain, and preferably present, (A); (e) complex of (A) or (I) with at least one additional molecule (III); in vitro detection of T cell activation by a compound containing at least one (A); (f) preparing cells of (d); and (g) test system for in vitro detection of T cells. The products of the invention have antiviral and antitumor activity. (A), also compounds or complexes containing it or nucleic acid or vectors encoding it, is used for detection of an immune response, particularly detecting L1-specific

CC cytotoxic T cells or to determine the L1 protein-specific antigenicity of
 CC compounds and complexes that contain (A), and for inducing a protective
 CC immune response against papilloma virus and related tumors, particularly
 CC as vaccines
 XX
 SQ Sequence 9 AA;

Query Match 100.0%; Score 55; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQLFNKPYW 9
 |||||
 DB 1 AQLFNKPYW 9

RESULT 2
 ABP47256
 ID ABP47256 standard; peptide; 20 AA.

AC ABP47256;
 XX
 DT 29-AUG-2003 (revised)
 DT 07-AUG-2003 (revised)
 DT 15-AUG-2002 (first entry)
 XX
 DE HPV16 L1 T-cell epitope peptide P28 SEQ ID NO 28.
 XX
 DE HPV; human papillomavirus; T-cell epitope; L1; E7; virucide; cytostatic;
 KM vaccine; immune response.
 XX

OS Human papillomavirus type 16.
 XX
 PN WO20024384-A2.
 XX
 PD 06-JUN-2002.
 XX
 PF 30-NOV-2001; 2001WO-EP014037.
 XX
 PR 01-DEC-2000; 2000DE-01059631.
 XX
 PA (MEDI-) MEDIGENE AG.
 XX
 PI Nieland J, Kaufmann A;
 XX
 DR WPI; 2002-471833/50.
 XX

PT New T cell epitopes derived from human papilloma virus, useful for
 PT detecting immune activation and in a vaccine against papilloma virus.
 XX
 PS Example 1; Page 35; 126pp; German.

XX The invention relates to T-cell epitopes (I) derived from human
 CC papillomavirus (HPV). (I) have virucide and cytostatic activity and are
 CC able to induce a cytotoxic T cell response or mediate T helper cell
 CC function. (I) and also compounds (II) containing them, vectors that
 CC express (I) or (II), cells that contain (I) or (II) and/or complexes of
 CC (I) or (II) with another compound, are used to detect or induce an immune
 CC response, for diagnosis or therapy, particularly in vaccines for control
 CC of human papilloma virus infection. The present sequence is that of a HPV
 CC T-cell epitope of the invention. (Updated on 07-AUG-2003 to correct OS
 CC field.) (Updated on 29-AUG-2003 to standardise OS field)
 XX
 SQ Sequence 20 AA;

Query Match 100.0%; Score 55; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.0056;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQLFNKPYW 9
 |||||
 DB 7 AQLFNKPYW 15

RESULT 3
 AAY71464
 ID AAY71464 standard; protein; 497 AA.

AC AAY71464;
 XX
 DT 04-OCT-2000 (first entry)
 XX
 DE Chimeric biotin-binding human papillomavirus mutant L1 protein-168.
 XX

XX Human papillomavirus; HPV; L1 protein; L2 protein; biotin; chimeric;
 KM cytostatic; antiviral; gene therapy; vaccine; capsomere; VLP;
 KM virus-like particle; viral antigen; tumour; cytotoxin; cytokine; IL;
 KW interleukin; chemotherapeutic agent; radioactive agent; mutant; mutein.
 XX

OS Human papillomavirus type 16.
 OS Unidentified.
 OS Chimeric.

PH Key Location/Qualifiers
 FT Region 1..483
 FT Binding-site /label= HPV_16_mutant_L1_protein
 FT /label=.497
 FT /label= Biotin_binding_peptide
 FT Region 484..486
 FT /note="(Gly)3 hinge"
 XX

PN WO200031128-A1.
 XX

PD 02-JUN-2000.
 XX

PF 22-NOV-1999; 99WO-US027555.
 XX

PR 23-NOV-1998; 98US-0109510P.
 PR 06-OCT-1999; 99US-00413611.
 XX

PA (LOYO) UNIV LOYOLA CHICAGO.
 XX

PI Mueller M, Kast WM, Nieland JD, Velders MP;
 XX

DR WPI; 2000-400041/34.
 DR N-PSDB; AAD01237.
 XX

PT Chimeric protein comprising a papillomavirus L1 or L2 protein and a
 PT biotin-binding polypeptide, useful for delivering substances such as
 PT proteins, nucleic acids and lipids into cells, particularly
 PT papillomavirus infected cells.
 XX

PS Example 1; Page 18-20; 27pp; English.

XX The patent discloses a chimeric protein comprising papillomavirus L1 or
 CC L2 protein and a biotin-binding polypeptide. Capsomere, papillomavirus or
 CC virus-like particle (VLP) comprising the chimeric protein is used for
 CC delivering a wide variety of biotinylated compounds e.g. proteins,
 CC nucleic acids and lipids into cells, particularly papillomavirus infected
 CC cells. The chimeric protein may also be used as a vaccine when the
 CC biotinylated substance is a viral antigen e.g. papillomavirus E2 or E7
 CC proteins. It may be useful for treating tumours or other papillomavirus-
 CC related lesions when the substance is a cytotoxin, chemotherapeutic
 CC agent, radioactive agent, or a gene encoding a cytokine or interleukin.
 CC The present sequence is a chimeric protein 168 consisting of human
 CC papillomavirus 16 (HPV-16) mutant L1 protein fused to a biotin-binding
 CC peptide at the C-terminus. The L1 protein was derived from the wild-type
 CC sequence by deleting the amino acid Cys428. The modification was done to
 CC prevent assembly of VLPs while allowing the production of capsomeres at
 CC high efficiencies
 XX

SQ Sequence 497 AA;

Query Match 100.0%; Score 55; DB 3; Length 497;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQLFNKPYW 9
 |||||
 304 AQLFNKPYW 312

Db

RESULT 4
 AAY71465
 ID AAY71465 standard; protein; 499 AA.
 XX AAY71465;
 AC
 XX 04-OCT-2000 (first entry)
 DT
 XX Chimeric biotin-binding human papillomavirus mutant L1 protein-169.
 DE
 XX Human papillomavirus; HPV; L1 protein; L2 protein; biotin; chimeric;
 KW cytosolic; antiviral; gene therapy; vaccine; capsomere; VLP;
 KM virus-like particle; viral antigen; tumour; cytotoxin; cytokine; IL;
 KM interleukin; chemotherapeutic agent; radioactive agent; mutant; mutein.
 XX
 OS Human papillomavirus type 16.
 OS unidentified.
 OS Chimeric.
 XX
 FH Key Location/Qualifiers
 FT Region 1. .483
 FT /label= HPV_16_mutant_L1_protein
 FT Binding-site 484. .499
 FT /label= Biotin_binding_domain
 FT 484. .486
 FT /note= "(Gly)3 hinge"
 FT
 XX WO200031128-A1.
 PN
 XX 02-JUN-2000.
 PD
 XX 22-NOV-1999; 99WO-US027555.
 PF
 XX 23-NOV-1998; 98US-0109510P.
 PR 06-OCT-1999; 99US-00413611.
 PR
 XX (LOYO) UNIV LOYOLA CHICAGO.
 PA
 XX Mueller M, Kast WM, Nieland JD, Velders MP;
 PI
 XX WPI: 2000-400041/34.
 DR N-PSDB; AAD01238.
 DR
 XX Chimeric protein comprising a papillomavirus L1 or L2 protein and a
 FT biotin-binding polypeptide, useful for delivering substances such as
 FT proteins, nucleic acids and lipids into cells, particularly
 FT papillomavirus infected cells.
 FT
 XX Example 1; Page 23-24; 27pp; English.
 PS
 XX The patent discloses a chimeric protein comprising papillomavirus L1 or
 CC L2 protein and a biotin-binding polypeptide. Capsomere, papillomavirus or
 CC virus-like particle (VLP) comprising the chimeric protein is used for
 CC delivering a wide variety of biotinylated compounds e.g. proteins,
 CC nucleic acids and lipids into cells, particularly papillomavirus infected
 CC cells. The chimeric protein may also be used as a vaccine when the
 CC biotinylated substance is a viral antigen e.g. papillomavirus E2 or E7
 CC proteins. It may be useful for treating tumours or other papillomavirus-
 CC related lesions when the substance is a cytotoxin, chemotherapeutic
 CC agent, radioactive agent, or a gene encoding a cytokine or interleukin.
 CC The present sequence is a chimeric protein 169 consisting of human
 CC papillomavirus 16 (HPV-16) mutant L1 protein fused to a biotin-binding
 CC peptide at the C-terminus. The L1 protein was derived from the wild-type
 CC sequence by deleting the amino acid Cys428. The modification was done to
 CC prevent assembly of VLPs while allowing the production of capsomeres at
 CC high efficiencies

SO Sequence 499 AA;
 Query Match 100.0%; Score 55; DB 3; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQLFNKPYW 9
 |||||
 304 AQLFNKPYW 312

Db

RESULT 5
 AAB98436
 ID AAB98436 standard; protein; 504 AA.
 XX AAB98436;
 AC
 XX 22-AUG-2001 (first entry)
 DT
 XX Human papillomavirus protein HPV31 L1.
 DE
 XX Human papillomavirus; human leukocyte antigen; HLA; immune response; HPV;
 KW epitope; T cell; identification; vaccine; infection; genital wart;
 KM neoplastic growth; antiviral.
 XX
 OS Human papillomavirus.
 OS
 PN WO200141799-A1.
 PD
 XX 14-JUN-2001.
 PD
 XX 11-DEC-2000; 2000WO-US03549.
 PF
 XX 10-DEC-1999; 99US-0172705P.
 PR 15-AUG-2000; 2000US-00641528.
 PR
 XX (EPIM-) EPIMMUNE INC.
 PA
 XX Sette A, Sidney J, Southwood S, Chesnut R, Celis E, Grey HM;
 PI
 XX WPI: 2001-381497/40.
 DR
 XX An isolated human papilloma virus (HPV) epitope, useful in vaccines for
 FT treating HPV infections.
 FT
 XX Disclosure; Page 23-24; 756pp; English.
 PS
 XX The present invention describes an isolated prepared human papillomavirus
 CC (HPV) epitope (I). (I) has antiviral activity, and can be used in vaccine
 CC production. Peptides and corresponding nucleic acid compositions from the
 CC present invention are useful for stimulating an immune response to HPV by
 CC stimulating the production of CTL or HTL responses, specifically in the
 CC treatment or prophylaxis of HPV infection, in persons who have not
 CC manifested symptoms e.g. genital warts or neoplastic growth. The peptides
 CC can also be used in a tetramer staining assay to assess peripheral blood
 CC mononuclear cells for the presence of antigen-specific CTLs following
 CC exposure to a pathogen or immunogen, and as reagents to evaluate immune
 CC recall responses or evaluate the efficacy of a vaccine. The vaccine
 CC compositions are useful for removing warts or treating HPV infections.
 CC The epitopes for inclusion in an epitope-base vaccine may be selected
 CC from conserved regions of viral or tumour-associated antigens, which
 CC reduces the likelihood of escape mutants, also immunosuppressive epitopes
 CC that may be present in whole antigens can be avoided with the use of
 CC epitope-base vaccines. An additional advantage is the ability to combine
 CC selected epitopes (CTL and HTL) and to modify the composition of the
 CC epitopes achieving enhanced immunogenicity, the major benefit of the
 CC vaccine is that its safe and efficacious. AAB98391 to AAB98477 represent
 CC polypeptide sequences used in the exemplification of the present
 CC invention
 XX
 XX Sequence 504 AA;
 Query Match 100.0%; Score 55; DB 4; Length 504;


```

FT      Protein                               1.,.50S  

FT      /label= L1  

PT      /note= "Partial capsid protein"  

XX  

XX  

XX      US5716620-A.  

XX  

XX      10-FEB-1998.  

XX  

XX      07-JUN-1995;       95US-00475783.  

XX      .  

XX      03-SEP-1992;       92US-00941371.  

PR      16-MAR-1993;       93US-00032869.  

BR  

XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.  

PA  

XX      Kirnbauer R, Lowy DR, Schiller JT;  

PI  

XX      WPl, 1998-158363/14.  

DR      N-PSDB; AAV09855.  

XX  

XX      Vaccine against human papilloma virus - comprises HPV16 L1 polypeptide.  

PT  

XX      Disclosure; Col 17-20; 20pp; English.  

XX      .  

XX  

XX      This sequence represents the L1 capsid protein from Human Papillomavirus  

CC      strain 16 (HPV16). A recombinant form of this viral protein which is  

CC      capable of self-assembly into capsomer structures and viral capsids that  

CC      comprise conformational antigenic epitopes can be used as a vaccine for  

CC      the prevention or treatment of papillomavirus infections in vertebrates.  

CC      The vaccine comprises a unit dose of a composition containing a self  

CC      assembled HPV16 with at least 1 papillomavirus L1 conformational epitope  

SQ      Sequence 505 AA;  

  

Query Match               100.0%; Score 55; DB 2; Length 505;  

Best Local Similarity    100.0%; Pred. No. 0.15;  

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0.  

OY      1 AQIFNKPYW 9  

        |||||  

Db      304 AQIFNKPYW 312  

  

RESULT 9  

AAW63116  

ID      AAW63116 standard; protein; 505 AA.  

XX  

XX      AAW63116;  

AC      XX  

DT      08-OCT-1998 (first entry)  

XX  

DE      Human papillomavirus type 16 VLP protein D202B mutant.  

DE  

XX      Human papillomavirus type 16; HPV16; virus-like particle; VLP;  

XX      HPV16D202B; H16_U4 antibody; H16_V5 antibody; identification;  

KW      HPV16 type; characterisation; HPV type 16 infection.  

KV      XX  

OS      Synthetic.  

OS      Human papillomavirus.  

OS  

FH      Key Location/Qualifiers  

FH      Misc-difference 202  

FT      /label= D202B  

FT      XX  

XX      WO9825646-A1.  

PD      18-JUN-1998.  

XX  

PF      05-DEC-1997;       97MO-US022023.  

XX  

FR      09-DEC-1996;       96US-0032633P.  

XX  

PA      (MERI ) MERCK & CO INC.
```

XX	ludmerer S;
PI	
XX	WPI, 1998-348266/30.
DR	
XX	
XX	Human papillomavirus type 16 D202E virus-like particles - used for
PT	characterising HPV16 type infections.
PS	Claim 1; Page; 37pp; English.
XX	
CC	The present sequence represents a synthetic human papillomavirus type 16
CC	(HPV16) virus-like particle (VLP) which contains a Glu substitution for
CC	the naturally occurring Asp at position 202. The protein is designated
CC	HPV16D202E. The novel VLP HPV:D202E binds H16.04 antibody but not H16.V5
CC	antibody and thus allows specific identification of HPV types. The new
CC	VLPs are used in the characterisation of HPV type 16 infections. note:
CC	the sequence does not appear in the specification. It was created using
CC	the sequence given in Genbank accession number AF084952, and information
CC	provided in the specification
XX	
SQ	Sequence 505 AA;
XX	
Query Match	100.0%; Score 55; DB 2; Length 505;
Best Local Similarity	100.0%; Pred. No. 0.15; Mismatches 0; Gaps 0;
Matches	9; Conservative 0; Indels 0;
OY	1 AQIFNKPYW 9
Db	304 AQIFNKPYW 312
XX	
RESULT 10	
AAW53487	
ID	AAW53487 standard; protein; 505 AA.
XX	
AC	AAW53487;
XX	
DT	10-JUN-1998 (first entry)
XX	
DE	Human papillomavirus proto-type HPV 16 L1 genome 5637-7155 protein.
XX	
XX	Human papillomavirus; HPV 16; L1 gene; immunisation; capsid;
KM	conformational epitope; vaccine; sequeae; vertebrate.
XX	
OS	Human papillomavirus.
XX	
PN	US5744142-A.
XX	
PD	28-APR-1998.
XX	
PF	07-JUN-1995; 95US-00475782.
XX	
PR	03-SEP-1992; 92US-00941371.
XX	
PR	16-MAR-1993; 93US-00032869.
XX	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX	
PI	Schiller JT, Kirmbauer R, Lowy DR;
XX	
DR	WPI; 1998-271010/24.
XX	
DR	N-PSDB; AAV25818.
XX	
PT	Immunisation of mammals and humans against papillomavirus infection -
PT	comprises administering recombinant self-assembled capsid proteins
PT	containing conformational epitopes.
XX	
PS	Disclosure; Col 17-20; 20pp; English.
XX	
CC	The present sequence represents the human papillomavirus proto-type HPV
CC	16 L1 genome 5637-7155 proteins, from the present invention. The present
CC	invention describes the immunisation of a host mammal against a
CC	papillomavirus comprising administering to the host, according to an
CC	immunising schedule, a self-assembled Human Papilloma Virus (HPV16)

CC capsid (I) containing at least 1 papillomavirus L1 conformational
 CC epitope. The papillomavirus L1 conformational epitopes are produced by
 CC letting a genetic construct comprising a papillomavirus L1 gene direct
 CC recombinant expression of the conformational epitope in a transformed
 CC eukaryotic host cell by self-assembly of papillomavirus capsids
 CC containing a L1 polypeptide having an amino acid sequence encoded by the
 CC nucleic acid sequence given in the specification (see AAV25817). (I) are
 CC useful for the diagnosis of and as vaccines for the prevention of
 CC papillomavirus infections and their benign and malignant sequelae in
 CC vertebrates. Recombinant conformed papillomavirus proteins are provided
 CC which can be used to produce renewable papillomavirus reagents of any
 CC selected species and type in cell culture. The self-assembled recombinant
 CC L1 capsid protein has the efficacy of intact papillomavirus particles to
 CC induce high levels of neutralising antiserum, in contrast to prior art L1
 CC protein extracted from recombinant bacteria or denatured virions

Query Match 100.0%; Score 55; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15; Mismatches 0; Gaps 0;
 Matches 9; Conservative 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
 |||||
 Db 304 AQIFNKPYW 312

RESULT 11
 AAM53486
 ID AAM53486 standard; protein; 505 AA.
 XX
 AC AAM53486;

DT 10-JUL-1998 (first entry)

DE Human papillomavirus wild-type HPV 16 L1 conformational epitope.

KM Human papillomavirus; HPV 16; L1 gene; immunisation; capsid;
 KW conformational epitope; vaccine; sequelae; vertebrate.

OS Human papillomavirus.

PN US5744142-A.

PD 28-APR-1998.

PF 07-JUN-1995; 95US-00475782.

PR 03-SEP-1992; 92US-00941371.

PR 16-MAR-1993; 93US-00032869.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Schiller JT, Kirnbauer R, Lowy DR;

DR WPI, 1998-271010/24.

DR N-PSDB; AAV25817.

PT Immunisation of mammals and humans against papillomavirus infection -
 PT comprises administering recombinant self-assembled capsid proteins
 PT containing conformational epitopes.

PS Claim 1; Col 21-24; 20pp; English.

XX The present sequence represents the human papillomavirus HPV 16 L1
 CC conformational epitope, from the present invention. The present invention
 CC describes the immunisation of a host mammal against a papillomavirus
 CC comprising administering to the host, according to an immunising
 CC schedule, a self-assembled Human Papilloma Virus (HPV16) capsid (I)
 CC containing at least 1 papillomavirus L1 conformational epitope. The
 CC papillomavirus L1 conformational epitopes are produced by letting a
 CC genetic construct comprising a papillomavirus L1 gene direct recombinant
 CC expression of the conformational epitope in a transformed eukaryotic host

CC cell by self-assembly of papillomavirus capsids containing a L1
 CC polypeptide having an amino acid sequence encoded by the nucleic acid
 CC sequence given in the specification (see AAV25817). (I) are useful for
 CC the diagnosis of and as vaccines for the prevention of papillomavirus
 CC infections and their benign and malignant sequelae in vertebrates.
 CC Recombinant conformed papillomavirus proteins are provided which can be
 CC used to produce renewable papillomavirus reagents of any selected species
 CC and type in cell culture. The self-assembled recombinant L1 capsid
 CC protein has the efficacy of intact papillomavirus particles to induce
 CC high levels of neutralising antiserum, in contrast to prior art L1
 CC protein extracted from recombinant bacteria or denatured virions

Query Match 100.0%; Score 55; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15; Mismatches 0; Gaps 0;
 Matches 9; Conservative 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
 |||||
 Db 304 AQIFNKPYW 312

RESULT 12
 AAM44142
 ID AAM44142 standard; protein; 505 AA.
 XX
 AC AAM44142;

DT 05-MAY-1998 (first entry)

DE Human papillomavirus wild-type L1 protein.

KM Human papillomavirus; bovine; L1 conformational protein; antibody;
 KW detection; diagnosis.

OS Human papillomavirus.

PN US5709996-A.

PD 20-JAN-1998.

PF 07-JUN-1995; 95US-00472673.

PR 03-SEP-1992; 92US-00941371.

PR 16-MAR-1993; 93US-00032869.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Schiller JT, Kirnbauer R, Lowy DR;

DR WPI, 1998-109811/10.

DR N-PSDB; AAV12161.

PT Assay for papilloma virus - using antibody directed against recombinant
 PT L1 polypeptide.

PS Claim 1; Col 21-24; 34pp; English.

XX An method has been developed for detecting papilloma virus in a specimen
 CC from a mammal. The method comprises contacting the sample with an
 CC antibody and detecting any antibody binding. The antibody is a polyclonal
 CC or monoclonal antibody produced by transforming a eukaryotic host cell
 CC with a genetic construct containing a papillomavirus L1 gene so that the
 CC cell expresses self-assembled papillomavirus-like particles containing at
 CC least 1 papillomavirus L1 conformational epitope. The present sequence
 CC represents wild-type human papillomavirus L1 protein. The method is for
 CC the diagnosis of papillomavirus infections in mammals

QY Sequence 505 AA;

Query Match 100.0%; Score 55; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
|||
304 AQIFNKPYW 312

RESULT 13

AAW4143
ID AAW4143 standard; protein; 505 AA.

AC AAW4143;

DT 05-MAY-1998 (first entry)

DE Human papillomavirus prototype L1 protein.

KW Human; papillomavirus; bovine; L1 conformational protein; antibody;

XX detection; diagnosis.

OS Human papillomavirus.

PN US5709996-A.

PD 20-JAN-1998.

PF 07-JUN-1995; 95US-00472673.

PR 03-SEP-1992; 92US-00941371.

PR 16-MAR-1993; 93US-00032869.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Schiller JT, Kirnbauer R, Lowy DR;

DR WPI: 1998-109811/10.

DR N-PSDB; AAV12162.

PT Assay for papilloma virus - using antibody directed against recombinant

PS L1 polypeptide.

PS Disclosure; Col 17-20; 34pp; English.

CC An method has been developed for detecting papilloma virus in a specimen
CC from a mammal. The method comprises contacting the sample with an
CC antibody and detecting any antibody binding. The antibody is a polyclonal
CC or monoclonal antibody produced by transforming a eukaryotic host cell
CC with a genetic construct containing a papillomavirus L1 gene so that the
CC cell expresses self-assembled papillomavirus-like particles containing at
CC least 1 papillomavirus L1 conformational epitope. The present sequence
CC represents prototype human papillomavirus L1 protein. The method is for
CC the diagnosis of papillomavirus infections in mammals

XX Sequence 505 AA;

Query Match 100.0%; Score 55; DB 2; Length 505;

Best Local Similarity 100.0%; Pred. No. 0.15; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
|||
304 AQIFNKPYW 312

RESULT 14

AAW54034
ID AAW54034 standard; protein; 505 AA.

AC AAW54034;

DT 17-OCT-2003 (revised)

DT 10-AUG-1998 (first entry)

DE HPV16 L1 gene protein #1.

KW Human papillomavirus 16; HPV17; L1 gene; infection; antibody detection.

OS Human papillomavirus type 16.

OS US5756284-A.

PD 26-MAY-1998.

PF 07-JUN-1995; 95US-00472672.

PR 03-SEP-1992; 92US-00941371.

PR 16-MAR-1993; 93US-00032869.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Kirnbauer R, Lowy DR, Schiller JT;

DR WPI: 1998-321522/28.

DR N-PSDB; AAV23923.

PT Assay for antibodies to papilloma virus - using self-assembled

PS papillomavirus-like particles comprising L1 polypeptide.

PS Disclosure; Col 17-20; 21pp; English.

CC This sequence represents the human papillomavirus 16 (HPV16) L1 gene
CC protein. The protein can be used in the method of the invention for
CC detecting antibodies to a papilloma virus in a sample from a mammal. The
CC method comprises: (a) providing self-assembled papillomavirus-like
CC particles comprising a L1 polypeptide, where the self-assembled
CC papillomavirus-like particles comprise at least one L1 conformational
CC epitope and are produced by permitting a genetic construct comprising a
CC papillomavirus L1 gene to direct recombinant expression in a transformed
CC eukaryotic host cell; (b) contacting the sample with the self-assembled
CC papillomavirus-like particles; and (c) detecting any antibody binding to
CC the self-assembled papillomavirus-like particles. The method can be used
CC for detecting antibodies to HPV16 as an indication of infection. (Updated
CC on 17-OCT-2003 to standardise OS field)

XX Sequence 505 AA;

Query Match 100.0%; Score 55; DB 2; Length 505;

Best Local Similarity 100.0%; Pred. No. 0.15; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
|||
304 AQIFNKPYW 312

RESULT 15

AAW54073
ID AAW54073 standard; protein; 505 AA.

AC AAW54073;

DT 17-OCT-2003 (revised)

DT 10-AUG-1998 (first entry)

DE HPV16 L1 gene protein #2.

KW Human papillomavirus 16; HPV17; L1 gene; infection; antibody detection.

OS Human papillomavirus type 16.

OS US5756284-A.

PD 26-MAY-1998.

PF 07-JUN-1995; 95US-00472672.

PR 03-SEP-1992; 92US-00941371.
 PR 16-MAR-1993; 93US-00032869.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Kirmbauer R, Lowy DR, Schiller JT;
 DR WPI, 1998-321522/28.
 DR N-PSDB; AAV24099.
 XX
 PT Assay for antibodies to papilloma virus - using self-assembled
 PT papillomavirus-like particles comprising L1 polypeptide.
 XX
 PS Claim 1; Col 21-24; 21pp; English.
 XX
 CC This sequence represents the human papillomavirus 16 (HPV16) L1 gene
 CC protein. The protein can be used in the method of the invention for
 CC detecting antibodies to a papilloma virus in a sample from a mammal. The
 CC method comprises: (a) providing self-assembled papillomavirus-like
 CC particles comprising a L1 polypeptide, where the self-assembled
 CC papillomavirus-like particles comprise at least one L1 conformational
 CC epitope and are produced by permitting a genetic construct comprising a
 CC papillomavirus L1 gene to direct recombinant expression in a transformed
 CC eukaryotic host cell; (b) contacting the sample with the self-assembled
 CC papillomavirus-like particles; and (c) detecting any antibody binding to
 CC the self-assembled papillomavirus-like particles. The method can be used
 CC for detecting antibodies to HPV16 as an indication of infection. (Updated
 CC on 17-Oct-2003 to standardise OS field)
 CC
 SQ Sequence 505 AA;
 XX
 Query Match 100.0%; Score 55; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQTENKPYW 9
 |||||
 DB 304 AQTENKPYW 312
 XX
 RESULT 16
 ID AAM96747 standard; protein; 505 AA.
 XX
 AC AAM96747;
 XX
 DT 14-APR-1999 (first entry)
 XX
 DE Human papillomavirus (HPV16) L1 major capsid protein fragment.
 XX
 KM HPV16; L1 major capsid protein; HPV16 capsid; L1 epitope;
 KM subunit vaccine; HPV infection; wart; laryngeal papillomatosis.
 XX
 OS Human papillomavirus.
 XX
 PN US5871998-A.
 XX
 PD 16-FEB-1999.
 XX
 PF 07-JUN-1995; 95US-00472678.
 XX
 PR 03-SEP-1992; 92US-00941371.
 PR 16-MAR-1993; 93US-00032869.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Kirmbauer R, Lowy DR, Schiller JT;
 XX
 DR WPI, 1999-166654/14.
 DR N-PSDB; AAX15111.
 XX
 PT Self-assembled, recombinant human papilloma virus 16 capsids - for use in
 PT subunit vaccines and as immunoassay reagents.

XX
 PS Disclosure; Col 17-20; 20pp; English.
 XX
 CC The present sequence represents a fragment of the prototype Human
 CC papillomavirus (HPV16) L1 major capsid protein. The sequence is used to
 CC construct a HPV16 capsid, containing at least one L1 conformational
 CC epitope. The capsids are produced by self-assembly following recombinant
 CC expression of a gene construct, containing an L1 gene, in a transformed
 CC eukaryotic cell. The capsids are used in subunit vaccines to generate a
 CC high-titre neutralizing antibody response in vertebrates, for treatment
 CC and prevention of HPV infection and its benign or malignant consequences,
 CC e.g. warts and laryngeal papillomatosis, including vaccination of
 CC pregnant women for passive protection of infants through the placenta or
 CC milk; as immunoassay reagents for diagnosing HPV infection; as reagents
 CC to detect cellular or humoral immunity, and as immunogens to generate
 CC antibodies (useful as immunoassay reagents for detecting capsid proteins)
 CC
 SQ Sequence 505 AA;
 XX
 Query Match 100.0%; Score 55; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQTENKPYW 9
 |||||
 DB 304 AQTENKPYW 312
 XX
 RESULT 17
 ID AAM96748 standard; protein; 505 AA.
 XX
 AC AAM96748;
 XX
 DT 14-APR-1999 (first entry)
 XX
 DE Human papillomavirus (HPV16) L1 major capsid protein fragment.
 XX
 KM HPV16; L1 major capsid protein; HPV16 capsid; L1 epitope;
 KM subunit vaccine; HPV infection; wart; laryngeal papillomatosis.
 XX
 OS Human papillomavirus.
 XX
 PN US5871998-A.
 XX
 PD 16-FEB-1999.
 XX
 PF 07-JUN-1995; 95US-00472678.
 XX
 PR 03-SEP-1992; 92US-00941371.
 PR 16-MAR-1993; 93US-00032869.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Kirmbauer R, Lowy DR, Schiller JT;
 XX
 DR WPI, 1999-166654/14.
 DR N-PSDB; AAX15112.
 XX
 PT Self-assembled, recombinant human papilloma virus 16 capsids - for use in
 PT subunit vaccines and as immunoassay reagents.
 XX
 PS Disclosure; Col 19-24; 20pp; English.
 XX
 CC The present sequence represents a fragment of the Human papillomavirus
 CC (HPV16) L1 major capsid protein. The sequence is used to construct a
 CC HPV16 capsid, containing at least one L1 conformational epitope. The
 CC capsids are produced by self-assembly following recombinant expression of
 CC a gene construct, containing an L1 gene, in a transformed eukaryotic
 CC cell. The capsids are used in subunit vaccines to generate a high-titre
 CC neutralizing antibody response in vertebrates, for treatment and
 CC prevention of HPV infection and its benign or malignant consequences,
 CC e.g. warts and laryngeal papillomatosis, including vaccination of

CC pregnant women for passive protection of infants through the placenta or
 CC milk; as immunoassay reagents for diagnosing HPV infection; as reagents
 CC to detect cellular or humoral immunity, and as immunogens to generate
 CC antibodies (useful as immunoassay reagents for detecting capsid proteins)
 XX
 SQ Sequence 505 AA;

Query Match 100.0%; Score 55; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
 |||||
 304 AQIFNKPYW 312

RESULT 18
 ID AAY08019 standard; protein; 505 AA.
 XX
 AC AAY08019;

DT 27-AUG-2003 (revised)
 DT 08-JUL-1999 (first entry)

DE Human papilloma virus L1 protein.

XX L1 protein; capsomer; virus; vaccine; infection; treatment; prevention;
 XX cervical carcinoma; fusion protein; anti-capsid; antibody; antigenicity.

XX Human papillomavirus.

XX WO918220-A1.

XX 15-APR-1999.

XX 06-OCT-1998; 98WO-US020965.

XX 06-OCT-1997; 97US-00944368.

XX (LOYO) UNIV LOYOLA CHICAGO.

XX Giesmann L, Mueller M;

XX WPI; 1999-264026/22.

XX DR N-PSDB; AAX37566.

XX Human papilloma virus (HPV) L1 fusion protein capsomers, used in vaccines
 PT against HPV infection.

XX Disclosure; Page 31-32; 48pp; English.

XX This invention describes novel vaccines comprising a human papilloma
 CC virus (HPV) capsomer, composed of a HPV L1 fusion protein or a HPV L1
 CC truncated protein. The vaccines can prevent or treat human papilloma
 CC virus infection. Such therapeutic vaccinations can be used for relief
 CC against, e.g. cervical carcinoma. Construction of chimeric proteins
 CC comprising amino acid residues from L1 protein, and e.g. B6 or B7
 CC protein, which give rise to chimeric capsomers, combines prophylactic and
 CC therapeutic functions of a vaccine. Capsomers can promote elimination of
 CC persistently infected cells. Capsomers can also escape neutralization by
 CC pre-existing anti-capsid antibodies and hence possess longer circulating
 CC half-life as compared to chimeric virus-like particles. The fusion
 CC protein, which forms the capsomer, provides increased antigenicity.
 CC (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 505 AA;

Query Match 100.0%; Score 55; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9

DB 304 AQIFNKPYW 312
 |||||

RESULT 19
 ID AAY53580 standard; protein; 505 AA.
 XX
 AC AAY53580;

DT 18-FEB-2000 (first entry)

DE HPV16 L1 protein.

XX Human papillomavirus; HPV16; L1; construct; major capsid protein; wart;
 XX recombinant; expression; epitope; antibody; capsomer; diagnosis; vaccine;
 XX antiserum; communicable papillomatosis; lesion; benign; infection;
 XX laryngeal papillomatosis.

XX Human papillomavirus.

XX US5985610-A.

XX 16-NOV-1999.

XX 07-JUN-1995; 95US-00484503.

XX 03-SEP-1992; 92US-00941371.

XX 16-MAR-1993; 93US-00032869.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Kirnbauer R, Schiller JT, Lowy DR;

XX WPI; 2000-012790/01.

XX DR N-PSDB; AAZ40564.

XX Genetic construct comprising a human papillomavirus gene is useful in
 PT diagnosis and prevention of papillomavirus infections.

XX Disclosure; Col 21-22; 21pp; English.

XX This sequence represents the human papillomavirus type 16 (HPV16) L1
 CC protein. The invention relates to a genetic construct comprising an HPV16
 CC gene for the L1 major capsid protein, which directs recombinant
 CC expression of L1 conformational epitopes used for raising neutralizing
 CC antibodies in a host cell by self-assembly of capsomer structures
 CC comprising a HPV16 L1 polypeptide. The construct is used for the
 CC diagnosis and prevention of papillomavirus infections. The recombinant L1
 CC papillomavirus capsid proteins are able to induce high titer neutralizing
 CC antiserum making them suitable for use against communicable
 CC papillomatosis. The vaccines can be used to treat productive
 CC laryngeal papillomatosis

XX Sequence 505 AA;

Query Match 100.0%; Score 55; DB 3; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
 |||||
 304 AQIFNKPYW 312

RESULT 20

ID AAY53579 standard; protein; 505 AA.

XX AAY53579;

DT 18-FEB-2000 (first entry)

XX HPV16 L1 protein - corrected version.
 DE
 XX
 XX Human papillomavirus; HPV16; L1; construct; major capsid protein; wart;
 KW recombinant; expression; epitope; antibody; capsomer; diagnosis; vaccine;
 KM anticapsid; communicable papillomatosis; lesion; benign; infection;
 KM laryngeal papillomatosis.
 XX
 OS Human papillomavirus.
 XX
 XX Key Location/Qualifiers
 FH
 FT Misc-difference 202 /note= "replaces Asp residue in wild type sequence given
 FT in the GenBank Accession No: K02718"
 FT Misc-difference 423 /note= "insertion of this residue as compared to the wild
 FT type sequence given in the GenBank Accession No: K02718"
 FT Misc-difference 439..440 /note= "deletion of an Asp residue from between these
 FT residues as compared to the wild type sequence given in
 FT the GenBank Accession No: K02718"
 FT
 XX US985610-A.
 PN
 XX 16-NOV-1999.
 PD
 XX 07-JUN-1995; 95US-00484503.
 PF
 XX 03-SEP-1992; 92US-00941371.
 PR
 PR 16-MAR-1993; 93US-00032869.
 PT
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Kirnbauer R, Schiller JT, Lowy DR;
 PI
 XX WPI; 2000-012790/01.
 DR
 DR N-PSDB; AA240563.
 XX
 PT Genetic construct comprising a human papillomavirus gene is useful in
 PT diagnosis and prevention of papillomavirus infections.
 XX
 XX
 PS Disclosure; Col 17-20; 21pp; English.
 XX
 CC This sequence represents a corrected version of the human papillomavirus
 CC type 16 (HPV16) L1 protein as compared to the HPV16 L1 gene sequence
 CC given in the prototype sequence (GenBank Accession No: K02718). The
 CC invention relates to a genetic construct comprising an HPV16 gene for the
 CC L1 major capsid protein, which directs recombinant expression of L1
 CC conformational epitopes used for raising neutralizing antibodies in a
 CC host cell by self-assembly of capsomer structures comprising a HPV16 L1
 CC polypeptide. The construct is used for the diagnosis and prevention of
 CC papillomavirus infections. The recombinant L1 papillomavirus capsid
 CC proteins are able to induce high titer neutralizing antiserum making them
 CC suitable for use against communicable papillomatosis. The vaccines can be
 CC used to treat productive papillomavirus lesions that occur in benign
 CC infections such as warts or laryngeal papillomatosis
 XX
 SQ Sequence 505 AA;
 Query Match 100.0%; Score 55; DB 3; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQTFNKPYW 9
 DB 304 AQTFNKPYW 312
 AC
 AC AAY57720;
 AC AAY57720 standard; protein; 505 AA.
 ID AAY57720
 XX
 AC AAY57720;

XX 14-MAR-2000 (first entry)
 DT
 XX
 XX Human papillomavirus L1 fusion protein SEQ ID NO:2.
 DE
 XX Human papillomavirus; HPV; L1 fusion protein; vaccine; cytostatic;
 KW viral capsomere; virucide; dermatological; malignant tumour formation;
 KM cervical cancer; cervical intraepithelial neoplasia; genital wart;
 KM condylomata acuminata.
 XX
 OS Human papillomavirus.
 OS Synthetic.
 OS
 XX CA2229955-A1.
 PN
 XX 20-AUG-1999.
 PD
 XX 20-FEB-1998; 98CA-02229955.
 PF
 XX 20-FEB-1998; 98CA-02229955.
 PR
 PR 20-FEB-1998; 98CA-02229955.
 XX
 PA (MEDT-) MEDIGENE GMBH.
 XX
 XX Burger A, Hallek M;
 PI
 XX WPI; 2000-063092/06.
 DR
 DR N-PSDB; AA248174.
 XX
 PT Fusion proteins comprising papillomavirus specific proteins useful for
 PT vaccinating against malignant tumors of the anogenital tract such as
 PT cervical carcinomas.
 XX
 XX
 PS Claim 5; Page 32-33; 46pp; English.
 XX
 CC The present sequence represents the specifically claimed fusion protein
 CC designated L1, which comprises 2 amino acids sequences from 2 different
 CC papillomavirus specific (PVS) proteins. The fusion protein may be
 CC administered for preventing and treating papillomavirus infections in
 CC humans and animals. Papillomaviruses are implicated in the pathology of
 CC malignant tumour formation in the anogenital tract (of these tumours,
 CC cervical cancer is the most frequent (50000 cases/year) and in the
 CC formation of precursor lesions of cervical intraepithelial neoplasia
 CC (CIN). Papillomaviruses also cause benign genital warts such as
 CC condylomata acuminata. However, the type and severity of disease caused
 CC by the papillomavirus is dependent on the strain causing the infection
 XX
 SQ Sequence 505 AA;
 Query Match 100.0%; Score 55; DB 3; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQTFNKPYW 9
 DB 304 AQTFNKPYW 312
 AC
 AC ABB77479;
 AC ABB77479 standard; protein; 505 AA.
 ID ABB77479
 XX
 AC ABB77479;
 AC ABB77479;
 DT 22-UTL-2002 (first entry)
 XX
 XX HPV16-L1 2.
 DE
 XX Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;
 KM HPV16-L1; cytostatic; virucide.
 XX
 OS Human papillomavirus.
 OS
 XX WO200238769-A2.
 PN

XX 16-MAY-2002.
PD 19-SEP-2001; 2001WO-DE003618.
XX 09-NOV-2000; 2000DE-01055545.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA (IPRP-) IPRK INST PFLANZENGENETIK & KULTURPFLANZE.
XX Mueller M, Leder C, Kleinschmidt J, Somewald U, Biemelt S;
PI WPI; 2002-426950/45.
DR N-PSDB; ABL58977, ABL58982.
XX New DNA sequences encoding human papilloma virus L1 or L2 protein, useful
PT in vaccines, are optimized for high-level expression in eukaryotic cells.
XX Claim 1; Fig 6; 39pp; German.
XX The invention relates to DNA sequences (I) that encode human papilloma
CC virus (HPV) 16 L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)
CC or proteins with the biological activity of L1 and L2. Expression vectors
CC containing (I) or a similar sequence encoding an L1/E7 fusion protein
CC (ABL58979-ABL58981) and the proteins encoded by them (ABB77478-ABB77483),
CC are useful in vaccines, especially for control of cervical cancer. (I)
CC are also useful for recombinant production of L1 and L2 proteins. (I) are
CC optimised for codon usage in eukaryotic cells and provide high yields of
CC L1/L2 or their fusions, without the use of viral vectors
XX SQ Sequence 505 AA;
OY Query Match 100.0%; Score 55; DB 5; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 AQIFNKPYW 9
304 AQIFNKPYW 312
RESULT 23
ABB77478
ID ABB77478 standard; protein; 505 AA.
XX ABB77478;
XX 22-JUL-2002 (first entry)
XX DT HPV16-L1 1.
XX DE HPV16-L1 1.
XX Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;
KW HPV16-L1; cytosolic; virucide.
XX OS Human papillomavirus.
XX PN WO200238769-A2.
XX 16-MAY-2002.
XX 19-SEP-2001; 2001WO-DE003618.
XX PR 09-NOV-2000; 2000DE-01055545.
XX PT (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA (IPRP-) IPRK INST PFLANZENGENETIK & KULTURPFLANZE.
XX Mueller M, Leder C, Kleinschmidt J, Somewald U, Biemelt S;
PI WPI; 2002-426950/45.
DR N-PSDB; ABL58976.
XX New DNA sequences encoding human papilloma virus L1 or L2 protein, useful
PT

PT in vaccines, are optimized for high-level expression in eukaryotic cells.
XX Claim 1; Fig 5; 39pp; German.
XX The invention relates to DNA sequences (I) that encode human papilloma
CC virus (HPV) 16 L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)
CC or proteins with the biological activity of L1 and L2. Expression vectors
CC containing (I) or a similar sequence encoding an L1/E7 fusion protein
CC (ABL58979-ABL58981) and the proteins encoded by them (ABB77478-ABB77483),
CC are useful in vaccines, especially for control of cervical cancer. (I)
CC are also useful for recombinant production of L1 and L2 proteins. (I) are
CC optimised for codon usage in eukaryotic cells and provide high yields of
CC L1/L2 or their fusions, without the use of viral vectors
XX SQ Sequence 505 AA;
OY Query Match 100.0%; Score 55; DB 5; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 AQIFNKPYW 9
304 AQIFNKPYW 312
RESULT 24
ABG75800
ID ABG75800 standard; protein; 505 AA.
XX ABG75800;
XX 23-OCT-2003 (revised)
XX DT 08-MAY-2003 (first entry)
XX DE Human papillomavirus 16 L1 protein, clone P14/16/11.
XX HPV16; virus-like particle; VLP; infection; cancer; tumour; L1;
KW anogenital tract; uterine cervix; clone P14/16/11; virucide; vaccine.
XX OS Human papillomavirus type 16.
XX PN US2002168372-A1.
XX 14-NOV-2002.
XX 29-SEP-1998; 98US-00162904.
XX PF 16-JUL-1993; 93US-00092528.
XX PR 01-MAY-1996; 96US-00641570.
XX PR 27-JUN-1997; 97US-00884168.
XX (DURS/) DURST M.
PA (GISS/) GISSMANN L.
XX Durst M, Gissmann L;
PI WPI; 2003-275319/27.
DR N-PSDB; ABX11396.
XX New DNA sequence encoding an L1 protein of a papillomavirus capable of
PT forming virus-like particles, useful for preparing a vaccine against
PT papillomavirus infection, particularly human papillomavirus (HPV)16
PT infection.
XX Disclosure; Fig 2; 17pp; English.
XX The invention discloses a DNA sequence encoding an L1 protein of a human
CC papillomavirus (HPV16) capable of forming virus-like particles (VLPs).
CC Papillomaviruses infections have been linked to malignant cancer and
CC malignant tumours of the anogenital tract, particularly cancer of the
CC uterine cervix. Also disclosed is an antibody that is specifically
CC directed against the VLP or the L1 protein, a method for determining anti-
CC -HPV16 virion antibodies in a sample and prophylaxis of papillomavirus

CC infections. The DNA sequence is useful for preparing a vaccine against
 CC papillomavirus infection, particularly HPV16 infection. The DNA is used
 CC to produce a VLP which can be used in a diagnostic kit to determine anti-
 CC HPV16 virion antibodies in a sample. The VLP can also be used for the
 CC prophylaxis of papillomavirus infections and the DNA used to produce an
 CC LI protein. The sequence presented is the HPV16 LI protein, clone
 CC P114/16/11. (Updated on 23-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 505 AA;
 Query Match 100.0%; Score 55; DB 6; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 A01FNKRPYW 9
 DB 304 A01FNKRPYW 312
 RESULT 25
 ABG75799
 ID ABG75799 standard; protein; 505 AA.
 XX
 AC ABG75799;
 XX
 DT 23-OCT-2003 (revised)
 DT 08-MAY-2003 (first entry)
 XX
 DE Human papillomavirus 16 LI protein, clone P114/16/2.
 XX
 KW HPV16; virus-like particle; VLP; infection; cancer; tumour; LI;
 KW anogenital tract; uterine cervix; clone P114/16/2; virucide; vaccine.
 XX
 OS Human papillomavirus type 16.
 XX
 PN US2002168372-A1.
 XX
 PD 14-NOV-2002.
 XX
 PF 29-SEP-1998; 98US-00162904.
 XX
 PR 16-JUL-1993; 93US-00092528.
 PR 01-MAY-1996; 96US-00641570.
 PR 27-JUN-1997; 97US-00864168.
 XX
 PA (DURS/) DURST M.
 PA (GISS/) GISSMANN L.
 PI Durst M, Gissmann L;
 XX
 DR WPI, 2003-275319/27.
 DR N-PSDB; ABX11395.
 XX
 PT New DNA sequence encoding an LI protein of a papillomavirus capable of
 PT forming virus-like particles, useful for preparing a vaccine against
 PT papillomavirus infection, particularly human papillomavirus (HPV) 16
 PT infection.
 PT
 PS Disclosure; Fig 1; 17pp; English.
 XX
 PS The invention discloses a DNA sequence encoding an LI protein of a human
 CC papillomavirus (HPV16) capable of forming virus-like particles (VLPS).
 CC Papillomaviruses infections have been linked to malignant cancer and
 CC malignant tumours of the anogenital tract, particularly cancer of the
 CC uterine cervix. Also disclosed is an antibody that is specifically
 CC directed against the VLP or the LI protein, a method for determining anti-
 CC HPV16 virion antibodies in a sample and prophylaxis of papillomavirus
 CC infections. The DNA sequence is useful for preparing a vaccine against
 CC papillomavirus infection, particularly HPV16 infection. The DNA is used
 CC to produce a VLP which can be used in a diagnostic kit to determine anti-
 CC HPV16 virion antibodies in a sample. The VLP can also be used for the
 CC prophylaxis of papillomavirus infections and the DNA used to produce an
 CC LI protein. The sequence presented is the HPV16 LI protein, clone

CC P114/16/2. (Updated on 23-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 505 AA;
 Query Match 100.0%; Score 55; DB 6; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 A01FNKRPYW 9
 DB 304 A01FNKRPYW 312
 RESULT 26
 ADA27363
 ID ADA27363 standard; protein; 505 AA.
 XX
 AC ADA27363;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE HPV-16 LI protein SEQ ID NO:6.
 XX
 KW cell line; American Type Culture Collection PTA-4047; ATCC-4047;
 KW baculoviruses; viral recombinant protein; virus-like particle; vaccine;
 KW diagnostic reagent; human papillomavirus type 16; HPV-16; LI.
 XX
 OS Human papillomavirus type 16.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 70
 FT /note= "encoded by GTC"
 FT Misc-difference 99
 FT /note= "encoded by TAC"
 XX
 PN WO2003068804-A2.
 XX
 PD 21-AUG-2003.
 XX
 PF 14-FEB-2003; 2003WO-US004516.
 XX
 PR 14-FEB-2002; 2002US-0356113P.
 PR 14-FEB-2002; 2002US-0356118P.
 PR 14-FEB-2002; 2002US-0356119P.
 PR 14-FEB-2002; 2002US-0356123P.
 PR 14-FEB-2002; 2002US-0356126P.
 PR 14-FEB-2002; 2002US-0356133P.
 PR 14-FEB-2002; 2002US-0356135P.
 PR 14-FEB-2002; 2002US-0356150P.
 PR 14-FEB-2002; 2002US-0356151P.
 PR 14-FEB-2002; 2002US-0356152P.
 PR 14-FEB-2002; 2002US-0356154P.
 PR 14-FEB-2002; 2002US-0356156P.
 PR 14-FEB-2002; 2002US-0356157P.
 PR 14-FEB-2002; 2002US-0356161P.
 PR 14-FEB-2002; 2002US-0356162P.
 XX
 PA (NOVA-) NOVAVAX INC.
 XX
 PI Robinson RA;
 XX
 DR WPI, 2003-646475/61.
 DR N-PSDB; ADA27371.
 XX
 PT New insect cell line designated ATCC PTA-4047, useful for replicating
 PT baculoviruses to produce large amounts of recombinant proteins of
 PT medical, pharmaceutical and veterinary importance.
 XX
 PS Disclosure; Page 55-56; 63pp; English.
 XX
 XX The present invention describes a cell line comprising a cell that is a
 CC clone, derivative, mutant and/or transfectant of a cell line designated
 CC American Type Culture Collection (ATCC) PTA-4047. The cell upon culture

CC grows continuously and retains the identifying characteristics of the
 CC cell line designated ATCC-4047. Also described is a process of making a
 CC cell line. The insect cell line is useful in replicating baculoviruses,
 CC as a host substrate for baculovirus plaque assays, as a source of insect
 CC proteins, acts as a depot for cell transfection to produce recombinant
 CC baculoviruses, and in expressing viral recombinant proteins.
 CC Extracellular and intracellular viral recombinant proteins and virus-like
 CC particles expressed from the cell line are useful as pharmaceutical
 CC compositions, vaccines or diagnostic reagents. The present sequence
 CC represents a human papillomavirus type 16 (HPV-16) L1 protein, which is
 CC used in the exemplification of the present invention.
 XX
 SQ Sequence 505 AA;
 QY
 Query Match 100.0%; Score 55; DB 6; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 304 A01FNKPYW 312
 RESULT 27
 ADA92541
 ID ADA92541 standard; protein; 505 AA.
 XX
 AC ADA92541;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE HPV-16 L1 codon optimised amino acid sequence SEQ ID NO:6.
 XX
 KW codon optimised; viral capsid protein; virus-like particle; VLP;
 KW antigenic; human papillomavirus infection; virucide; vaccine;
 KW gene therapy; human papillomavirus type 16; dysplasia; infection; HPV-16;
 L1.
 XX
 OS Synthetic.
 OS Human papillomavirus type 16.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 70
 FT Misc-difference 99 /note= "encoded by GTC"
 FT Misc-difference 99 /note= "encoded by TAC"
 XX
 PD WO2003068933-A2.
 XX
 PD 21-AUG-2003.
 XX
 PF 14-FEB-2003; 2003WO-US004480.
 XX
 PR 14-FEB-2002; 2002US-0356113P.
 PR 14-FEB-2002; 2002US-0356118P.
 PR 14-FEB-2002; 2002US-0356119P.
 PR 14-FEB-2002; 2002US-0356123P.
 PR 14-FEB-2002; 2002US-0356126P.
 PR 14-FEB-2002; 2002US-0356133P.
 PR 14-FEB-2002; 2002US-0356135P.
 PR 14-FEB-2002; 2002US-0356150P.
 PR 14-FEB-2002; 2002US-0356151P.
 PR 14-FEB-2002; 2002US-0356152P.
 PR 14-FEB-2002; 2002US-0356154P.
 PR 14-FEB-2002; 2002US-0356156P.
 PR 14-FEB-2002; 2002US-0356157P.
 PR 14-FEB-2002; 2002US-0356161P.
 PR 14-FEB-2002; 2002US-0356162P.
 XX
 XX (NOVA-) NOVAVAX INC.
 PA Robinson RA;
 XX

DR WPI; 2003-689664/65.
 DR N-PEDB; ADA92549.
 XX
 FT New codon optimised polynucleotide encoding a viral capsid protein that
 PT self assembles into a virus-like particle, useful for diagnosing,
 PT preventing or treating human papillomavirus infections or associated
 PT disorders.
 XX
 PS Disclosure; Page 115-116; 123pp; English.
 XX
 CC The present invention describes a codon optimised polynucleotide encoding
 CC a viral capsid protein that self assembles into a virus-like particle
 CC (VLP) that exhibits conformational antigenic epitopes capable of raising
 CC neutralising antibodies, where the VLP is expressed from a host cell
 CC extracellularly. Also described: (1) a vector comprising the above codon
 CC optimised polynucleotide operably linked to a eukaryotic or prokaryotic
 CC regulatory control element, capable of replication in prokaryotic and/or
 CC eukaryotic host; (2) a host cell comprising the vector; (3) a
 CC pharmaceutical or vaccine composition for treating, ameliorating or
 CC preventing a papillomavirus related disease or disorder, comprising a
 CC multiplicity of VLPs that exhibit conformational antigenic epitopes, and
 CC a carrier, diluent or adjuvant; (4) a diagnostic kit for detecting a
 CC papillomavirus infection, comprising a multiplicity of VLPs that exhibit
 CC conformational antigenic epitopes, and a detection agent comprising a
 CC detectable label; (5) a method for preparing the above codon optimised
 CC polynucleotide, comprising replacing codons that are underutilised in
 CC insect cells with codons that are utilised at high levels in insect
 CC cells, to create an initially-modified nucleotide sequence, and modifying
 CC the initially-modified nucleotide sequence by choosing a preferred codon
 CC for the initially-modified sequence, where the ratio of GC nucleotide
 CC pairs to AT nucleotide pairs in the further-modified nucleotide sequence
 CC trends towards about 1:1, where the number of palindromic and stem-loop
 CC DNA structures in the further-modified nucleotide sequence is minimised,
 CC and where the number of transcription and post-transcription repressor
 CC elements are minimised; and (6) methods for treating, ameliorating or
 CC preventing a papillomavirus related disease or disorder, or for
 CC protecting an individual against a papillomavirus infection, comprising
 CC administering to an individual an amount of the composition or vaccine
 CC cited above. The VLP has virucide activity and can be used in vaccines
 CC and in gene therapy. The composition and methods of the present invention
 CC are useful in diagnosing, preventing or treating human papillomavirus
 CC infections or associated disorders, such as dysplasia. The present
 CC sequence represents an HPV-16 codon optimised L1 amino acid sequence from
 CC the present invention.
 XX
 SQ Sequence 505 AA;
 QY
 Query Match 100.0%; Score 55; DB 7; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 304 A01FNKPYW 312
 RESULT 28
 ADA14290
 ID ADA14290 standard; protein; 505 AA.
 XX
 AC ADA14290;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE HPV-16 L1 protein SEQ ID NO:6.
 XX
 KW purification; recombinant extracellular virus-like particle;
 KW recombinant intracellular virus-like particle; virus-like particle; VLP;
 KW virucide; vaccine; gene therapy; human papillomavirus; HPV; infection;
 KW dysplasia; HPV-16; L1.
 XX
 OS Human papillomavirus type 16.
 XX

```

FH Key Location/Qualifiers
FT Misc-difference 70
FT /note= "encoded by GTC"
FT Misc-difference 99
FT /note= "encoded by TAC"
XX
XX WO2003068993-A1.
XX
XX 21-AUG-2003.
XX
XX 14-FEB-2003; 2003WO-US004474.
XX
XX 14-FEB-2002; 2002US-0356113P.
XX 14-FEB-2002; 2002US-0356118P.
XX 14-FEB-2002; 2002US-0356119P.
XX 14-FEB-2002; 2002US-0356123P.
XX 14-FEB-2002; 2002US-0356126P.
XX 14-FEB-2002; 2002US-0356133P.
XX 14-FEB-2002; 2002US-0356135P.
XX 14-FEB-2002; 2002US-0356150P.
XX 14-FEB-2002; 2002US-0356151P.
XX 14-FEB-2002; 2002US-0356152P.
XX 14-FEB-2002; 2002US-0356154P.
XX 14-FEB-2002; 2002US-0356156P.
XX 14-FEB-2002; 2002US-0356157P.
XX 14-FEB-2002; 2002US-0356161P.
XX 14-FEB-2002; 2002US-0356162P.
XX
XX (NOVA-) NOVAVAX INC.
XX
XX Robinson RA, Thompson MW,
XX
XX WPI, 2003-679645/64.
XX
XX N-PSDB; ADA14298.
XX
XX Purifying a recombinant human papillomavirus (HPV) L1, useful for
XX diagnosing, preventing or treating HPV infections, comprises clarifying,
XX concentrating and dialyzing cells containing HPV particles.
XX
XX Disclosure; Page 99-100; 11pp; English.
XX
XX The present invention describes a method for purifying a recombinant
XX extracellular or intracellular virus-like particle (VLP). The method
XX comprises harvesting a cell suspension comprising cells containing a
XX plurality of VLPS to produce a harvested supernatant, optionally
XX disrupting the harvested cells to produce cell lysates containing the
XX VLP, clarifying the harvested supernatant, concentrating the clarified
XX supernatant, dialyzing the concentrated supernatant, and recovering
XX the purified recombinant VLP. Also described: (1) a cell line designated
XX as sf-9s deposited as American Type Culture Collection (ATCC) PTA-4047;
XX (2) producing the cell line described above; (3) host cells that express
XX one or more recombinant gene products with an enhanced yield; (4)
XX producing a foreign protein in an insect cell; (5) nucleic acid sequences
XX that correspond to and code for human papillomavirus (HPV) polypeptides;
XX and (6) pharmaceutical compositions comprising an amount of the
XX recombinant viral gene products, VLPS, agonists, antagonists, or the
XX active fragment of a viral gene product. The VLPS have virucide activity,
XX and can be used in vaccines and in gene therapy. The method is useful in
XX isolating and purifying expressed viral gene products, including VLPS, in
XX vitro. The gene products or particles may be used in detecting,
XX preventing or treating HPV infections and associated symptoms like
XX dysplasia. The present sequence represents an HPV-16 L1 protein sequence,
XX which is given in the exemplification of the present invention.
XX
XX Sequence 505 AA;
XX
XX Query Match 100.0%; Score 55; DB 7; Length 505;
XX Best Local Similarity 100.0%; Pred. No. 0.15;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

XX
XX RESULT 29
XX AAE38615
XX ID AAE38615 standard; protein; 505 AA.
XX
XX AAE38615;
XX
XX 04-DEC-2003 (first entry)
XX
XX HPV-16 L1 protein.
XX
XX Vaccine; humoral immunity; cell-mediated immunity; gene therapy; HPV;
XX virucide; papillomavirus infection; Human papillomavirus.
XX
XX Human papillomavirus.
XX
XX Key Location/Qualifiers
XX
XX Misc-difference 70
XX /note= "Encoded by GTC"
XX
XX Misc-difference 99
XX /note= "Encoded by TAC"
XX
XX WO2003068163-A2.
XX
XX 21-AUG-2003.
XX
XX 14-FEB-2003; 2003WO-US004473.
XX
XX 14-FEB-2002; 2002US-0356113P.
XX 14-FEB-2002; 2002US-0356118P.
XX 14-FEB-2002; 2002US-0356119P.
XX 14-FEB-2002; 2002US-0356123P.
XX 14-FEB-2002; 2002US-0356126P.
XX 14-FEB-2002; 2002US-0356133P.
XX 14-FEB-2002; 2002US-0356135P.
XX 14-FEB-2002; 2002US-0356150P.
XX 14-FEB-2002; 2002US-0356151P.
XX 14-FEB-2002; 2002US-0356152P.
XX 14-FEB-2002; 2002US-0356154P.
XX 14-FEB-2002; 2002US-0356156P.
XX 14-FEB-2002; 2002US-0356157P.
XX 14-FEB-2002; 2002US-0356161P.
XX 14-FEB-2002; 2002US-0356162P.
XX
XX (NOVA-) NOVAVAX INC.
XX
XX Robinson RA, Cioce V,
XX
XX WPI, 2003-689598/65.
XX
XX N-PSDB; AAD58571.
XX
XX New chimeric virus-like particles comprising a recombinant viral capsid
XX protein encapsulating a recombinant viral protein, useful for inducing
XX humoral and/or cell-mediated immunity against papillomavirus infection.
XX
XX Disclosure; Page 118-119; 126pp; English.
XX
XX The present invention relates to chimeric virus-like particle comprising
XX a recombinant viral capsid protein that encapsulates a recombinant viral
XX protein during self assembly into a chimeric virus-like particle and
XX exhibiting conformational antigenic epitopes capable of eliciting
XX neutralizing antibodies. The vaccine comprising the chimeric virus-like
XX particles are useful for inducing immunity (humoral and/or cell-mediated
XX immunity) against papillomavirus infection. The invention is also useful
XX in gene therapy. The present sequence is HPV (human papillomavirus)-16 L1
XX protein
XX
XX Sequence 505 AA;
XX
XX Query Match 100.0%; Score 55; DB 7; Length 505;
XX Best Local Similarity 100.0%; Pred. No. 0.15;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


QY 1 AQIFNKPYW 9
 DB 304 AQIFNKPYW 312

RESULT 30
 ID ADC26174 standard; protein, 505 AA.
 XX
 AC ADC26174;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE HPV16 L1 major capsid protein.
 XX
 KW HPV16; virucide; vaccine; laryngeal papillomatosis; genital wart;
 KW gene therapy; L1 major capsid protein.
 XX
 OS Human papillomavirus.
 XX
 US2003050439-A1.
 PD 13-MAR-2003.
 PF 09-APR-2001; 2001US-00832065.
 XX
 PR 03-SEP-1992; 92US-00941371.
 PR 16-MAR-1993; 93US-00032869.
 PR 07-JUN-1995; 95US-00484503.
 PR 21-MAY-1999; 99US-00316487.
 XX
 PA (LOWY/) LOWY D R.
 PA (SCHI/) SCHILLER J T.
 PA (KIRN/) KIRNBAUER R.
 PI Lowy DR, Schiller JT, Kirnbauer R;
 DR WPI; 2003-615802/58.
 DR N-PSDB; ADC26168.
 XX
 PT New HPV16 L1 polypeptide, useful for manufacturing a medicament for
 PT diagnosing, treating or preventing papillomavirus infection.
 XX
 PS Disclosure; Page 10-11; 18pp; English.
 CC The invention relates to a novel human papillomavirus (HPV)16 L1
 CC polypeptide, assembled in a capsomeric structure possessing
 CC conformational epitopes and made by expression of a recombinant coding
 CC sequence in a human. The polypeptide of the invention demonstrates
 CC virucide activity and may be useful for manufacturing a medicament or
 CC vaccine for diagnosing, treating or preventing papillomavirus infection,
 CC such as laryngeal papillomatosis and genital warts. The polypeptide may
 CC also be utilised for gene therapy purposes. The current sequence is that
 CC of the HPV16 L1 major capsid protein of the invention.
 XX
 SQ Sequence 505 AA;

Query Match 100.0%; Score 55; DB 7; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
 DB 304 AQIFNKPYW 312

RESULT 31
 ID ADC26175 standard; protein, 505 AA.
 XX
 AC ADC26175;
 XX

DT 18-DEC-2003 (first entry)
 XX
 DE HPV16 L1 major capsid protein-related protein.
 XX
 KW HPV16; virucide; vaccine; laryngeal papillomatosis; genital wart;
 KW gene therapy; L1 major capsid protein; human papillomavirus.
 XX
 OS Unidentified.
 XX
 US2003050439-A1.
 PD 13-MAR-2003.
 PF 09-APR-2001; 2001US-00832065.
 XX
 PR 03-SEP-1992; 92US-00941371.
 PR 16-MAR-1993; 93US-00032869.
 PR 07-JUN-1995; 95US-00484503.
 PR 21-MAY-1999; 99US-00316487.
 XX
 PA (LOWY/) LOWY D R.
 PA (SCHI/) SCHILLER J T.
 PA (KIRN/) KIRNBAUER R.
 PI Lowy DR, Schiller JT, Kirnbauer R;
 DR WPI; 2003-615802/58.
 DR N-PSDB; ADC26169.
 XX
 PT New HPV16 L1 polypeptide, useful for manufacturing a medicament for
 PT diagnosing, treating or preventing papillomavirus infection.
 XX
 PS Disclosure; Page 12-13; 18pp; English.
 CC The invention relates to a novel human papillomavirus (HPV)16 L1
 CC polypeptide, assembled in a capsomeric structure possessing
 CC conformational epitopes and made by expression of a recombinant coding
 CC sequence in a human. The polypeptide of the invention demonstrates
 CC virucide activity and may be useful for manufacturing a medicament or
 CC vaccine for diagnosing, treating or preventing papillomavirus infection,
 CC such as laryngeal papillomatosis and genital warts. The polypeptide may
 CC also be utilised for gene therapy purposes. The current sequence is that
 CC of the HPV16 L1 major capsid protein-related protein of the invention.
 XX
 SQ Sequence 505 AA;

Query Match 100.0%; Score 55; DB 7; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
 DB 304 AQIFNKPYW 312

RESULT 32
 ID AAB23925 standard; protein, 531 AA.
 XX
 AC AAB23925;
 XX
 DT 18-JAN-2001 (first entry)
 XX
 DE Human papillomavirus 16 L1 protein SEQ ID NO:1.
 XX
 KW Human papillomavirus; HPV16; HPV18; L1 protein; L2 protein; antiviral;
 KW vaccine; immunisation; immune response; infection; diagnosis.
 XX
 OS Human papillomavirus.
 XX
 WO200054730-A2.
 PD 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US006017.
 PF 18-MAR-1999; 99US-0125208P.
 PR 12-AUG-1999; 99US-0148544P.
 XX
 PA (HARD) HARVARD COLLEGE.
 PA (HARR/) HARRISON S.
 PA (CHEN/) CHEN X.
 PI Harrison S., Chen X;
 XX WPI; 2000-628165/60.
 DR
 XX Composition comprising multimer of human papillomavirus L1 protein useful
 PT as vaccines against human papillomavirus infection, as diagnostic tools
 PT for detecting presence of the virus in biological samples.
 XX
 PS Claim 38; Page 30; 31pp; English.
 XX
 CC The present invention describes a composition (I) comprising a multimer
 CC of human papillomavirus (HPV) L1 protein (II) which is truncated at its
 CC amino terminus. (I) has antiviral activity and can be used in the
 CC production of a vaccine. (II) comprising a multimer of HPV L1 protein (a
 CC T-1 icosahedral particle comprising a pentamer of L1 protein, one of
 CC which comprises HPV L2 protein) when administered induces an immune
 CC response against the L1 protein in the humans and is thus useful for
 CC immunising the humans against HPV. (I) is thus useful as vaccines against
 CC human papillomavirus infection, as diagnostic tools for detecting the
 CC presence of HPV in biological samples and as tools for mapping receptor
 CC interactions. The present sequence represents an HPV16 L1 protein
 CC sequence which is used in the exemplification of the present invention.
 CC
 SQ Sequence 531 AA;
 XX
 Query Match 100.0%; Score 55; DB 3; Length 531;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 A0IFNKPW 9
 Db 330 A0IFNKPW 338
 XX
 RESULT 33
 ID AAB98422 standard; protein; 531 AA.
 AC AAB98422;
 XX
 DT 22-AUG-2001 (first entry)
 XX
 DE Human papillomavirus protein HPV16 L1.
 XX
 KW Human papillomavirus; human leukocyte antigen; HLA; immune response; HPV;
 KW epitope; T cell; identification; vaccine; infection; genital wart;
 KW neoplastic growth; antiviral.
 XX
 OS Human papillomavirus.
 XX
 PN WO200141799-A1.
 PD 14-JUN-2001.
 XX
 PF 11-DEC-2000; 2000WO-US033549.
 XX
 PR 10-DEC-1999; 99US-0172705P.
 PR 15-AUG-2000; 2000US-00641528.
 XX
 PA (BPIM-) EPIMUNE INC.
 PA
 PI Sette A, Sidney J, Southwood S, Cheenut R, Celis E, Grey HM;
 XX

DR WPI; 2001-381497/40.
 XX
 XX An isolated human papilloma virus (HPV) epitope, useful in vaccines for
 PT treating HPV infections.
 PT
 PS Disclosure; Page 21; 756pp; English.
 XX
 CC The present invention describes an isolated prepared human papillomavirus
 CC (HPV) epitope (I). (I) has antiviral activity, and can be used in vaccine
 CC production. Peptides and corresponding nucleic acid compositions from the
 CC present invention are useful for stimulating an immune response to HPV by
 CC stimulating the production of CTL or HTL responses, specifically in the
 CC treatment or prophylaxis of HPV infection, in persons who have not
 CC manifested symptoms e.g. genital warts or neoplastic growth. The peptides
 CC can also be used in a tetramer staining assay to assess peripheral blood
 CC mononuclear cells for the presence of antigen-specific CTLs following
 CC exposure to a pathogen or immunogen, and as reagents to evaluate immune
 CC recall responses or evaluate the efficacy of a vaccine. The vaccine
 CC compositions are useful for removing warts or treating HPV infections.
 CC The epitopes for inclusion in an epitope-base vaccine may be selected
 CC from conserved regions of viral or tumour-associated antigens, which
 CC reduces the likelihood of escape mutants, also immunosuppressive epitopes
 CC that may be present in whole antigens can be avoided with the use of
 CC epitope-base vaccines. An additional advantage is the ability to combine
 CC selected epitopes (CTL and HTL) and to modify the composition of the
 CC epitopes achieving enhanced immunogenicity. The major benefit of the
 CC vaccine is that is safe and efficacious. AAB98391 to AAB98477 represent
 CC polypeptide sequences used in the exemplification of the present
 CC invention
 CC
 SQ Sequence 531 AA;
 XX
 Query Match 100.0%; Score 55; DB 4; Length 531;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 A0IFNKPW 9
 Db 330 A0IFNKPW 338
 XX
 RESULT 34
 ID ABB77482
 AB B77482 standard; protein; 533 AA.
 AC ABB77482;
 XX
 DT 22-JUL-2002 (first entry)
 XX
 DE HPV16-L1/E7 fusion protein (short).
 XX
 KW Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;
 KW HPV16-L1; cytostatic; virucide; HPV16-E7; HPV16-L1/E7.
 XX
 OS Human papillomavirus.
 XX
 FH Key Location/Qualifiers
 FT Protein 1..533
 FT /product= "HPV16-L1/E7 fusion protein"
 FT Region 1..473
 FT /product= "HPV16-L1"
 FT Region 474..533
 FT /product= "HPV16-E7"
 XX
 PN WO200238769-A2.
 PD 16-MAY-2002.
 XX
 PF 19-SEP-2001; 2001WO-DE003618.
 XX
 PR 09-NOV-2000; 2000DE-01055545.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX

PA (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
XX
XX MueUler M, Leder C, Kleinschmidt J, Sonnwald U, Biemelt S;
PI
XX WPI; 2002-426950/45.
DR
XX N-PSDB; ABL58980.
DR
XX
XX New DNA sequences encoding human papilloma virus L1 or L2 protein, useful
PT in vaccines, are optimized for high-level expression in eukaryotic cells.
XX
XX Claim 4; Fig 9; 39pp; German.
XX
XX The invention relates to DNA sequences (I) that encode human papilloma
CC virus (HPV) 16 L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)
CC or proteins with the biological activity of L1 and L2. Expression vectors
CC containing (I) or a similar sequence encoding an L1/E7 fusion protein
CC (ABL58979-ABL58981) and the proteins encoded by them (AB877478-AB877483),
CC are useful in vaccines, especially for control of cervical cancer. (I)
CC are also useful for recombinant production of L1 and L2 proteins. (I) are
CC optimized for codon usage in eukaryotic cells and provide high yields of
CC L1/L2 or their fusions, without the use of viral vectors
CC
XX
XX Sequence 533 AA;
SQ
Query Match 100.0%; Score 55; DB 5; Length 533;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AQLFNKPYW 9
DB 304 AQLFNKPYW 312
RESULT 35
AB877481
ID AB877481 standard; protein; 533 AA.
XX
XX AB877481;
AC
XX
XX 22-JUL-2002 (first entry)
DT
XX
XX HPV16-L1/E7 fusion protein 1-60.
DE
XX
XX Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;
KW HPV16-L1; cytostatic; virucide; HPV16-E7; HPV16-L1/E7.
XX
XX Human papillomavirus.
OS
XX
XX Key Location/Qualifiers
FH 1..533
FT /product= "HPV16-L1/E7 fusion protein"
FT 1..473
FT /product= "HPV16-L1"
FT 474..533
FT /product= "HPV16-E7"
FT
XX
XX WO200238769-A2.
XX
XX 16-MAY-2002.
PD
XX
XX 19-SEP-2001; 2001WO-DE003618.
PF
XX
XX 09-NOV-2000; 2000DE-01055545.
PR
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
XX
XX MueUler M, Leder C, Kleinschmidt J, Sonnwald U, Biemelt S;
PI
XX
XX WPI; 2002-426950/45.
DR
XX
XX N-PSDB; ABL58979.
DR
XX
XX New DNA sequences encoding human papilloma virus L1 or L2 protein, useful
PT

PT in vaccines, are optimized for high-level expression in eukaryotic cells.
XX
XX Claim 4; Fig 8; 39pp; German.
XX
XX The invention relates to DNA sequences (I) that encode human papilloma
CC virus (HPV) 16 L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)
CC or proteins with the biological activity of L1 and L2. Expression vectors
CC containing (I) or a similar sequence encoding an L1/E7 fusion protein
CC (ABL58979-ABL58981) and the proteins encoded by them (AB877478-AB877483),
CC are useful in vaccines, especially for control of cervical cancer. (I)
CC are also useful for recombinant production of L1 and L2 proteins. (I) are
CC optimized for codon usage in eukaryotic cells and provide high yields of
CC L1/L2 or their fusions, without the use of viral vectors
CC
XX
XX Sequence 533 AA;
SQ
Query Match 100.0%; Score 55; DB 5; Length 533;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AQLFNKPYW 9
DB 304 AQLFNKPYW 312
RESULT 36
AB877483
ID AB877483 standard; protein; 542 AA.
XX
XX AB877483;
AC
XX
XX 22-JUL-2002 (first entry)
DT
XX
XX HPV16-L1/E7 fusion protein (long).
DE
XX
XX Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;
KW HPV16-L1; cytostatic; virucide; HPV16-E7; HPV16-L1/E7.
XX
XX Human papillomavirus.
OS
XX
XX Key Location/Qualifiers
FH 1..542
FT /product= "HPV16-L1/E7 fusion protein"
FT 1..473
FT /product= "HPV16-L1"
FT 474..542
FT /product= "HPV16-E7"
FT
XX
XX WO200238769-A2.
XX
XX 16-MAY-2002.
PD
XX
XX 19-SEP-2001; 2001WO-DE003618.
PF
XX
XX 09-NOV-2000; 2000DE-01055545.
PR
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
XX
XX MueUler M, Leder C, Kleinschmidt J, Sonnwald U, Biemelt S;
PI
XX
XX WPI; 2002-426950/45.
DR
XX
XX N-PSDB; ABL58981.
DR
XX
XX New DNA sequences encoding human papilloma virus L1 or L2 protein, useful
PT in vaccines, are optimized for high-level expression in eukaryotic cells.
XX
XX Claim 4; Fig 10; 39pp; German.
XX
XX The invention relates to DNA sequences (I) that encode human papilloma
CC virus (HPV) 16 L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)
CC or proteins with the biological activity of L1 and L2. Expression vectors
CC containing (I) or a similar sequence encoding an L1/E7 fusion protein
CC

CC (ABL58979-ABL58981) and the proteins encoded by them (AB77478-AB77483),
 CC are useful in vaccines, especially for control of cervical cancer. (1)
 CC are also useful for recombinant production of L1 and L2 proteins. (1) are
 CC optimised for codon usage in eukaryotic cells and provide high yields of
 CC L1/L2 or their fusions, without the use of viral vectors
 XX
 SQ Sequence 542 AA;

Query Match 100.0%; Score 55; DB 5; Length 542;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQLFNKPYW 9
 Db 304 AQLFNKPYW 312

RESULT 37
 AAW47281
 ID AAW47281 standard; protein; 128 AA.
 XX
 AC AAW47281;
 XX
 DT 25-MAR-2003 (revised)
 DT 30-JUL-1998 (first entry)
 XX
 DE Papillomavirus major capsid protein fragment from plasmid DL78.
 XX
 KM Major capsid protein; detection; infection; disease; wart; cancer; skin;
 KM diagnosis; treatment; prevention; vaccine; cervix; mouth; throat.
 XX
 OS Human papillomavirus.
 XX
 PN DE19648962-C1.
 XX
 PD 26-FEB-1998.
 XX
 PF 26-NOV-1996; 96DE-01048962.
 XX
 PR 26-NOV-1996; 96DE-01048962.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX
 PI De Villiers- Zur Hausen E, Zur Hausen H, Laverigne D, Benton C;
 DR WPI; 1998-121621/12.
 DR N-PSDB; AAV15599.
 XX
 PT DNA encoding peptide(s) from papilloma virus major capsid protein - and
 PT related proteins, vectors, transformed cells and antibodies, for
 PT diagnosis, treatment and prevention of papilloma virus infection and
 PT disease.
 XX
 PS Claim 1; Fig 6; 16pp; German.
 XX
 CC This sequence represents a fragment of the human papillomavirus (PV)
 CC major capsid protein isolated from plasmid DL78. This fragment could be
 CC used as a reagent for detecting PV infections and diseases (e.g. warts or
 CC cancers of skin, cervix or mouth/throat, particularly of the skin). Cells
 CC containing transformants can be used to produce recombinant
 CC papillomavirus which can be used for diagnosis, treatment and prevention
 CC (as vaccines) of papillomavirus infection. (Updated on 25-MAR-2003 to
 CC correct PI field.)
 CC
 SQ Sequence 128 AA;

Query Match 96.4%; Score 53; DB 2; Length 128;
 Best Local Similarity 88.9%; Pred. No. 0.084;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQLFNKPYW 9
 Db 110 AQLFNKPYW 118

RESULT 38
 AAW47280
 ID AAW47280 standard; protein; 128 AA.
 XX
 AC AAW47280;
 XX
 DT 25-MAR-2003 (revised)
 DT 30-JUL-1998 (first entry)
 XX
 DE Papillomavirus major capsid protein fragment from plasmid DL40.
 XX
 KM Major capsid protein; detection; infection; disease; wart; cancer; skin;
 KM diagnosis; treatment; prevention; vaccine; cervix; mouth; throat.
 XX
 OS Human papillomavirus.
 XX
 PN DE19648962-C1.
 XX
 PD 26-FEB-1998.
 XX
 PF 26-NOV-1996; 96DE-01048962.
 XX
 PR 26-NOV-1996; 96DE-01048962.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX
 PI De Villiers- Zur Hausen E, Zur Hausen H, Laverigne D, Benton C;
 DR WPI; 1998-121621/12.
 DR N-PSDB; AAV15599.
 XX
 PT DNA encoding peptide(s) from papilloma virus major capsid protein - and
 PT related proteins, vectors, transformed cells and antibodies, for
 PT diagnosis, treatment and prevention of papilloma virus infection and
 PT disease.
 XX
 PS Claim 1; Fig 5; 16pp; German.
 XX
 CC This sequence represents a fragment of the human papillomavirus (PV)
 CC major capsid protein isolated from plasmid DL40. This fragment could be
 CC used as a reagent for detecting PV infections and diseases (e.g. warts or
 CC cancers of skin, cervix or mouth/throat, particularly of the skin). Cells
 CC containing transformants can be used to produce recombinant
 CC papillomavirus which can be used for diagnosis, treatment and prevention
 CC (as vaccines) of papillomavirus infection. (Updated on 25-MAR-2003 to
 CC correct PI field.)
 CC
 SQ Sequence 128 AA;

Query Match 96.4%; Score 53; DB 2; Length 128;
 Best Local Similarity 88.9%; Pred. No. 0.084;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQLFNKPYW 9
 Db 110 AQLFNKPYW 118

RESULT 39
 AAO16091
 ID AAO16091 standard; protein; 467 AA.
 XX
 AC AAO16091;
 XX
 DT 27-FEB-2003 (first entry)
 DT Human papillomavirus truncated L1 protein.
 XX
 DE Gene therapy; vaccine; humoral immune response; cellular immune response;
 KW immune response modulation; pathogenic infection; rheumatoid arthritis.
 XX

```

OS Human papillomavirus type 6b.
XX
XX WO200283181-A1.
XX
XX 24-OCT-2002.
XX
XX 18-APR-2002; 2002WO-AU000486.
XX
XX 18-APR-2001; 2001AU-00004468.
XX
XX (UYOU ) UNIV QUEENSLAND.
XX
XX Frazer IH;
XX
XX WPI; 2003-075509/07.
XX
XX N-PSDB; AAL51109.
XX
XX New compositions having antigens, polynucleotides encoding the antigens,
XX or antigen-presenting cells, useful for modulating an immune response,
XX e.g. for treating or preventing pathogenic infections or rheumatoid
XX arthritis.
XX
XX Disclosure; Page 96-98; 139pp; English.
XX
XX The invention comprises compositions for eliciting a humoral or cellular
XX immune response against a target antigen. The compositions of the
XX invention are useful for eliciting a humoral and cellular immune response
XX against a target antigen, modulating an immune response in a patient, and
XX the treatment or prophylaxis of a disease or condition. This includes
XX pathogenic infections (e.g. viral, bacterial, fungal or protozoan), or an
XX immune response to an autoantigen (e.g. rheumatoid arthritis). The
XX present amino acid sequence represents a protein which was used in the
XX invention
XX
XX Sequence 467 AA;
XX
XX Query Match 96.4%; Score 53; DB 6; Length 467;
XX Best Local Similarity 88.9%; Pred. No. 0.31;
XX Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AQLFNKPYW 9
XX ||:|||||
XX 300 AQLFNKPYW 308
XX
XX RESULT 40
XX AA016092
XX ID AA016092 standard; protein; 467 AA.
XX
XX AA016092;
XX
XX 27-FEB-2003 (first entry)
XX
XX Humanised HPV truncated L1 protein.
XX
XX Gene therapy; vaccine; humoral immune response; cellular immune response;
XX immune response modulation; pathogenic infection; rheumatoid arthritis.
XX
XX Human papillomavirus type 6b.
XX
XX Synthetic.
XX
XX WO200283181-A1.
XX
XX 24-OCT-2002.
XX
XX 18-APR-2002; 2002WO-AU000486.
XX
XX 18-APR-2001; 2001AU-00004468.
XX
XX (UYOU ) UNIV QUEENSLAND.
XX
XX Frazer IH;
XX
XX

```

```

DR WPI; 2003-075509/07.
DR N-PSDB; AAL51110.
XX
XX New compositions having antigens, polynucleotides encoding the antigens,
XX or antigen-presenting cells, useful for modulating an immune response,
XX e.g. for treating or preventing pathogenic infections or rheumatoid
XX arthritis.
XX
XX Disclosure; Page 100-102; 139pp; English.
XX
XX The invention comprises compositions for eliciting a humoral or cellular
XX immune response against a target antigen. The compositions of the
XX invention are useful for eliciting a humoral and cellular immune response
XX against a target antigen, modulating an immune response in a patient, and
XX the treatment or prophylaxis of a disease or condition. This includes
XX pathogenic infections (e.g. viral, bacterial, fungal or protozoan), or an
XX immune response to an autoantigen (e.g. rheumatoid arthritis). The
XX present amino acid sequence represents a protein which was used in the
XX invention
XX
XX Sequence 467 AA;
XX
XX Query Match 96.4%; Score 53; DB 6; Length 467;
XX Best Local Similarity 88.9%; Pred. No. 0.31;
XX Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AQLFNKPYW 9
XX ||:|||||
XX 300 AQLFNKPYW 308
XX

```

Search completed: March 3, 2004, 10:22:47
 Job time : 52.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 09:47:18 ; Search time 10 Seconds
(without alignments)

46.863 Million cell updates/sec

Title: US-09-980-064-1
Perfect score: 55
Sequence: 1 AQIFNKPYW 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	502	VL1_HPV35	P27232 human papill
2	55	100.0	503	VL1_HPV29	P50792 human papill
3	55	100.0	503	VL1_HPV32	P36737 human papill
4	55	100.0	504	VL1_HPV31	P17388 human papill
5	55	100.0	528	VL1_HPV34	P36738 human papill
6	55	100.0	531	VL1_HPV16	P03101 human papill
7	53	96.4	499	VL1_HPV13	P02273 human papill
8	53	96.4	499	VL1_HPV53	P05113 human papill
9	53	96.4	500	VL1_HPV6A	P03100 human papill
10	53	96.4	501	VL1_HPV11	P04012 human papill
11	53	96.4	501	VL1_HPV11	P22163 rhesus papill
12	53	96.4	502	VL1_HPV28	P50791 human papill
13	53	96.4	502	VL1_HPV42	P27233 human papill
14	53	96.4	502	VL1_HPV1	P02274 pygmy chimp
15	53	96.4	503	VL1_HPV26	P36735 human papill
16	53	96.4	503	VL1_HPV6	P08961 human papill
17	53	96.4	508	VL1_HPV30	P02515 human papill
18	53	96.4	531	VL1_HPV10	P36743 human papill
19	53	96.4	534	VL1_HPV56	P36744 human papill
20	52	94.5	504	VL1_HPV54	P26536 human papill
21	51	92.7	497	VL1_HPV54	P50819 human papill
22	50	90.9	499	VL1_HPV33	P06416 human papill
23	50	90.9	504	VL1_HPV70	P50793 human papill
24	50	90.9	505	VL1_HPV39	P24838 human papill
25	50	90.9	505	VL1_HPV61	P50822 human papill
26	50	90.9	505	VL1_HPV68	P54669 human papill
27	50	90.9	505	VL1_HPVME	P27964 human papill
28	50	90.9	524	VL1_HPV58	P26535 human papill
29	50	90.9	529	VL1_HPV52	P05138 human papill
30	50	90.9	539	VL1_HPV45	P36794 human papill
31	50	90.9	568	VL1_HPV18	P06794 human papill
32	49	89.1	510	VL1_HPV2A	P25486 human papill
33	49	89.1	510	VL1_HPV57	P22162 human papill

34	49	89.1	532	1	VL1_HPV03	P36731 human papill
35	49	89.1	594	1	VL1_HPV27	P36736 human papill
36	48	87.3	495	1	VL1_BPVI	P03103 bovine papill
37	48	87.3	497	1	VL1_BPVI	P06458 bovine papill
38	48	87.3	507	1	VL1_HPV63	P07861 human papill
39	47	85.5	513	1	VL1_HPV48	P50817 human papill
40	47	85.5	515	1	VL1_HPV50	P50818 human papill
41	47	85.5	516	1	VL1_HPV04	P07860 human papill
42	47	85.5	516	1	VL1_HPV65	P07874 human papill
43	46	83.6	501	1	VL1_PAPVS	P11326 european el
44	46	83.6	503	1	VL1_COPV	P08928 canine oral
45	46	83.6	507	1	VL1_HPV09	P02480 human papill

ALIGNMENTS

RESULT 1
VL1_HPV35 STANDARD; PRT; 502 AA.
AC P27232;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 35.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10587;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Deline H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92124753; PubMed=1310198;
RA Martin J.E., Pontsler A.V., Rice S.M., McGraw K.A., Dubensky T.W.;
RT "The phylogenetic relationship and complete nucleotide sequence of human papillomavirus type 35";
RL Virology 186:770-776(1992).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74477; AAA52566.1; -;
DR EMBL; M74117; AAA46972.1; -;
DR PIR; G40824; P1W135.
DR PIR; S36526; S36526.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coat protein: late protein.
FT CONFLICT 94 95 AS -> CL (IN REF. 2).
FT CONFLICT 132 132 S -> L (IN REF. 2).
FT CONFLICT 140 140 G -> GNSG (IN REF. 2).
FT CONFLICT 482 482 K -> R (IN REF. 2).
SQ SEQUENCE 502 AA; 56148 MW; 68F6C43508F2267 CRC64;
Query Match 100.0%; Score 55; DB 1; Length 502;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 302 AOIFNKPYW 310

RESULT 2

VL1_HPV29 STANDARD; PRT; 503 AA.

AC P50782;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DE Major capsid protein L1.

GN L1.

OS Human papillomavirus type 29.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.

OX NCBI_TaxID=37112;

RP SEQUENCE FROM N.A.

RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 316-467 FROM N.A.

RX MEDLINE=95052821; PubMed=7963696;

RA Bernard H.U., Chan S.Y., Manos M.M., Ong C.K., Villa L.L.,

RA Delius H., Peyton C.L., Bauer H.M., Wheeler C.M.;

RT Identification and assessment of known and novel human

RT papillomaviruses by polymerase chain reaction amplification,

RT restriction fragment length polymorphisms, nucleotide sequence, and

RT phylogenetic algorithms.";

RL J. Infect. Dis. 170:1077-1085 (1994).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U1784; AAA79435.1; -

DR EMBL; U12503; AAA67247.1; -

DR InterPro; IPR002210; PV_capsid_L1.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00500; late_protein_L1; 1.

DR PRINTS; PR00865; HPVcapsidL1.

DR ProDom; PD00544; PV_capsid_L1; 1.

KW Coat protein; late protein.

SQ SEQUENCE 503 AA; 56281 MW; 8D2E36DAABD30847 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 503;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AOIFNKPYW 9

Db 304 AOIFNKPYW 312

RESULT 3

VL1_HPV32 STANDARD; PRT; 503 AA.

AC P36737;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DE Major capsid protein L1.

GN L1.

OS Human papillomavirus type 32.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.

OX NCBI_TaxID=10612;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94265501; PubMed=8205838;

RA Delius H., Hofmann B.;

RT "Primer-directed sequencing of human papillomavirus types.;"

RL Curr. Top. Microbiol. Immunol. 186:13-31 (1994).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X74475; CAA52554.1; -

DR PIR; S36514; S36514.

DR InterPro; IPR002210; PV_capsid_L1.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00500; late_protein_L1; 1.

DR PRINTS; PR00865; HPVcapsidL1.

DR ProDom; PD00544; PV_capsid_L1; 1.

KW Coat protein; late protein.

SQ SEQUENCE 503 AA; 56294 MW; 54C570C7C3E04292 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 503;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AOIFNKPYW 9

Db 304 AOIFNKPYW 312

RESULT 4

VL1_HPV31 STANDARD; PRT; 504 AA.

AC P17388;

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Major capsid protein L1.

GN L1.

OS Human papillomavirus type 31.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.

OX NCBI_TaxID=10585;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89299478; PubMed=2545036;

RA Goldborough M.D., Dislieste D., Temple G.F., Lorincz A.T.;

RT "Nucleotide sequence of human papillomavirus type 31: a cervical

RT neoplasia-associated virus.;"

RL Virology 171:306-311 (1989).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; J04353; AAA46956.1; -

DR PIR; G32444; P1WJ31.

DR InterPro; IPR002210; PV_capsid_L1.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00500; late_protein_L1; 1.

DR PRINTS; PR00865; HPVcapsidL1.

DR ProDom; PD00544; PV_capsid_L1; 1.

KW Coat protein; late protein.

SQ SEQUENCE 504 AA; 56352 MW; B45A306AB3AB9D2 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 504;

Best Local Similarity 100.0%; Pred. No. 0.006; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
Db 305 AQIFNKPYW 313

RESULT 5

VL1 HPV34 STANDARD; PRT; 528 AA.
AC P36738;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 34.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10613;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Holmann B.;
RT "Primer-directed sequencing of human papillomavirus types";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X74476; CAAS5260.1; -.
DR PIR; S36520; S36520.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00665; HPVcapsidL1.
DR PRODOM; PD000544; PV_capsid_L1; 1.
RW Coat protein; Late protein.
SQ SEQUENCE 528 AA; 59048 MW; A3F92CA480BA3703 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 528;
Best Local Similarity 100.0%; Pred. No. 0.0063;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
Db 327 AQIFNKPYW 335

RESULT 6
VL1 HPV16 STANDARD; PRT; 531 AA.
ID VL1 HPV16
AC P03101;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85246220; PubMed=2990099;
RA Seedorf K., Kramer G., Durst M., Roweckamp W.G.;
RT "Human papillomavirus type 16 DNA sequence.";

RL Virology 145:181-185 (1985).

RN [2]
RP SEQUENCE OF 328-371 FROM N.A.
RX MEDLINE=92407963; PubMed=1326639;
RA Chan S.Y., Bernard H.U., Ong C.K., Chan S.P., Birgit H., Delius H.;
RT "Phylogenetic analysis of 48 papillomavirus types and 26 subtypes and
RT variants: a showcase for the molecular evolution of DNA viruses.";
RL J. Virol. 66:5714-5725(1992).
CC -1- MISCELLANEOUS: HPV16, IN COMPARISON TO HPV TYPES 6 AND 11, IS MORE
CC OFTEN ASSOCIATED WITH MALIGNANT GENITAL CANCERS IN HUMANS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; K02718; AAA46943.1; -.
DR EMBL; M96285; AAA47024.1; -.
DR EMBL; A06331; CAA00546.1; -.
DR PIR; A03640; P1WMLH.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00665; HPVcapsidL1.
DR PRODOM; PD000544; PV_capsid_L1; 1.
RW Coat protein; Late protein.
SQ SEQUENCE 531 AA; 59554 MW; 5B3402587093B380 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 531;
Best Local Similarity 100.0%; Pred. No. 0.0063;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
Db 330 AQIFNKPYW 338

RESULT 7
VL1 HPV13 STANDARD; PRT; 499 AA.
ID VL1 HPV13
AC Q02273;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 13.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10573;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92391075; PubMed=1325697;
RA van Ranst M., Fuse A., Fiten P., Beuken E., Pfister H., Burk R.D.,
RA Opendakker G.;
RT "Human papillomavirus type 13 and pygmy chimpanzee papillomavirus
RT type 1: comparison of the genome organizations.";
RL Virology 190:587-596(1992).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X62843; CAA44654.1; -.
DR PIR; H42955; P1WML3.
DR InterPro; IPR002210; PV_capsid_L1.

DR InterPro: IPR008975; Viral_cap_coat.
 DR Pfam: PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 KW Coat protein; Late protein.
 SQ SEQUENCE 499 AA; 55773 MW; AA6D5F7FFD24759 CRC64;

Query Match 96.4%; Score 53; DB 1; Length 499;
 Best Local Similarity 88.9%; Pred. No. 0.014;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AOFENKPYW 9
 DB 298 AOLFENKPYW 306

RESULT 8
 ID_VL1_HPV53 STANDARD; PRT; 499 AA.
 AC 005113;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Major capsid protein L1.
 GN L1.
 OS Human papillomavirus type 53.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10619;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94255501; PubMed=8205838;
 RA Delius H., Hofmann B.;
 RT "Primer-directed sequencing of human papillomavirus types.";
 RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
 RN [2]
 RP SEQUENCE OF 300-343 FROM N.A.
 RX MEDLINE=92407963; PubMed=1326639;
 RA Chan S.Y., Bernard H.O., Ong C.K., Chan S.P., Birgit H., Delius H.;
 RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and
 RT variants: a showcase for the molecular evolution of DNA viruses";
 RL J. Virol. 66:5714-5725(1992).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X74482; CAA52595.1; -;
 CC EMBL; M96298; AAA47037.1; -;
 CC PIR; S36531; S36531.
 DR InterPro: IPR002210; PV_capsid_L1.
 DR InterPro: IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 KW Coat protein; Late protein.
 SQ SEQUENCE 499 AA; 55722 MW; 5221961A3PDD5A66 CRC64;

Query Match 96.4%; Score 53; DB 1; Length 499;
 Best Local Similarity 88.9%; Pred. No. 0.014;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AOFENKPYW 9
 DB 302 AOLFENKPYW 310

RESULT 9
 VL1_HPV6A

ID VL1_HPV6A STANDARD; PRT; 500 AA.
 AC P03100;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Major capsid protein L1.
 GN L1.
 OS Human papillomavirus type 6a, and
 OS Human papillomavirus type 6b.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=37122, 10600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX SPECIES=Human papillomavirus type 6a;
 RX MEDLINE=95297152; PubMed=7778283;
 RA Hofmann K.D., Cook J.C., Joyce J.G., Brown D.R., Schultz L.D.,
 RA George H.A., Rosolowsky M., Fife K.H., Jansen K.U.;
 RT "Sequence determination of human papillomavirus type 6a and assembly
 RT of virus-like particles in Saccharomyces cerevisiae";
 RL Virology 209:506-518(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX SPECIES=Human papillomavirus type 6b;
 RX MEDLINE=84131949; PubMed=6321162;
 RA Schwarz E., Durst M., Demankowski C., Lattemann O., Zech R.,
 RA Wolfeberger E., Suhai S., Zur Hausen H.;
 RT "DNA sequence and genome organization of genital human papillomavirus
 RT type 6b.";
 RL EMBO J. 2:2341-2348(1983).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; LA1216; AAA74218.1; -;
 CC EMBL; X00203; CAA25026.1; -;
 CC PIR; A03638; P1WL6.
 DR InterPro: IPR002210; PV_capsid_L1.
 DR InterPro: IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 KW Coat protein; Late protein.
 SQ SEQUENCE 500 AA; 55597 MW; B6168D0F76A287F7 CRC64;

Query Match 96.4%; Score 53; DB 1; Length 500;
 Best Local Similarity 88.9%; Pred. No. 0.014;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AOLFENKPYW 9
 DB 300 AOLFENKPYW 308

RESULT 10
 ID_VL1_HPV11 STANDARD; PRT; 501 AA.
 AC P04012;
 DT 23-OCT-1986 (Rel. 02, Created)
 DT 23-OCT-1986 (Rel. 02, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Major capsid protein L1.
 GN L1.
 OS Human papillomavirus type 11.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10580;
 RN [1]

```

RX  SEQUENCE FROM N.A.. PubMed=3008427;
RA  MEDLINE=66181601;
RA  Dattmann K., Schwarz E., Giesmann L., Zur Hausen H.;
RT  "The nucleotide sequence and genome organization of human papilloma
RL  virus type 11."
RL  Virology 151:124-130(1986).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M14119; AAA46935.1; -.
DR  PIR; A03639; P1ML1.
DR  InterPro; IPR002210; PV_capsid.L1.
DR  InterPro; IPR008975; Viral_cap.coat.
DR  Pfam; PF00500; late_protein.L1; 1.
DR  PRINTS; PR00865; HPV_CABSID.L1.
DR  ProDom; PD000544; PV_capsid.L1; 1.
KW  Papillomavirus; Late protein.
SQ  SEQUENCE 501 AA; 55835 MW; 88957D023CC182A7 CRC64;

Query Match      96.4%; Score 53; DB 1; Length 501;
Best Local Similarity 88.9%; Pred. No. 0.014;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  1 AQIFNKPYM 9
Db  301 AQLFNKPYM 309

RESULT 11
V11_RHPV1 STANDARD; PRT; 501 AA.
AC  P22I63;
DT  01-AUG-1991 (Rel. 19; Created)
DT  01-FEB-1996 (Rel. 33; Last sequence update)
DT  01-OCT-1996 (Rel. 34; Last annotation update)
GN  Major capsid protein L1.
NM  L1.
OS  Rheus papillomavirus type 1 (Rhpv 1).
OC  Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC  Papillomavirus.
OX  NCBI_TaxID=10570;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=91135018; PubMed=1847267;
RA  Ostrow R.S., Labresh K.V., Fates A.J.;
RT  "Characterization of the complete Rhpv 1 genomic sequence and an
RT  integration locus from a metastatic tumor."
RL  Virology 181:424-429(1991).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M60184; AAA79318.1; -.
DR  PIR; H38503; P1ML1.
DR  InterPro; IPR002210; PV_capsid.L1.
DR  InterPro; IPR008975; Viral_cap.coat.
DR  Pfam; PF00500; late_protein.L1; 1.
DR  PRINTS; PR00865; HPV_CABSID.L1.
DR  ProDom; PD000544; PV_capsid.L1; 1.
KW  Coat protein; Late protein.
SQ  SEQUENCE 501 AA; 55635 MW; 6E80BD8F130E440A CRC64;

```

```

Query Match          96.4%; Score 53; DB 1; Length 501;
Best Local Similarity 88.9%; Pred. No. 0.014;
Matches      8; Conservative      1; Mismatches      0; Indels      0; Gaps      0;

QY      1 AQTFFKPYW 9
      ||:|||||
      ||:|||||
DB      301 AQTFFKPYW 309

RESULT 12
VL1 HPV28      STANDARD;      PRT;      502 AA.
ID VL1 HPV28
AC P50791;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 28.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
CX NCBI_TaxID=37111;
FP [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 316-464 FROM N.A.
RA MEDLINE=95052821; PubMed=7963696;
RA Bernard H.U., Chan S.Y., Manor M.M., Ong C.K., Villa L.L.,
RA Delius H., Peyton C.L., Bauer H.M., Wheeler C.M.;
RT "Identification and assessment of known and novel human
RT papillomaviruses by polymerase chain reaction amplification,
RT restriction fragment length polymorphisms, nucleotide sequence, and
RT phylogenetic algorithms.";
RL J. Infect. Dis. 170:1077-1085(1994).

-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC
DR EMBL; U31783; AAA679428.1; -.
DR EMBL; U31502; AAA67246.1; -.
DR InterPro; IPR00210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; Late_protein_L1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
DR KX Coats protein; Late protein.
DR KW SEQUENCE 502 AA; 55902 MW; 38D6EADB12AF310A CRC64;

Query Match          96.4%; Score 53; DB 1; Length 502;
Best Local Similarity 88.9%; Pred. No. 0.014;
Matches      8; Conservative      1; Mismatches      0; Indels      0; Gaps      0;

QY      1 AQTFFKPYW 9
      ||:|||||
      ||:|||||
DB      304 AQTFFKPYW 312

RESULT 13
VL1 HPV42
ID VL1 HPV42      STANDARD;      PRT;      502 AA.
AC P27233;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.

```

```

OS Human papillomavirus type 42.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
NCBI_TaxID=10590;
RN SEQUENCE FROM N.A.
RP MEDLINE=92087479; PubMed=1309278;
RX Phillip W.; Honore N.; Sapp M.; Cole S.T.; Streeck R.E.;
RA "Human papillomavirus type 42: new sequences, conserved genome
organization.";
RL Virology 186:331-334(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; WJ3236; AAA47048.1; ALT_INIT.
DR PIR; G39451; P1WL42.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
DR Coat protein; late_protein.
SQ SEQUENCE 502 AA; 56165 MW; ED8083E650A1EA2E CRC64;

Query Match
Best Local Similarity 96.4%; Score 53; DB 1; Length 502;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQLFNKPYW 9
DB 304 AQLFNKPYW 312

RESULT 14
VL1_PCPV1 STANDARD; PRT; 502 AA.
AC 002274;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Pygmy chimpanzee papillomavirus type 1.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
NCBI_TaxID=10576;
RN SEQUENCE FROM N.A.
RP MEDLINE=92391075; PubMed=1325697;
RX van Ranst M.; Fuse A.; Filten P.; Beuken E.; Pfeister H.; Burk R.D.;
RA Opdenaker G.;
RT "Human papillomavirus type 13 and pygmy chimpanzee papillomavirus
RT type 1: comparison of the genome organizations.";
RL Virology 190:587-596(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X62844; CAA44662.1; -.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.

```

```

DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
DR Coat protein; late_protein.
SQ SEQUENCE 502 AA; 55718 MW; A22D1980EEFA94BA3 CRC64;

Query Match
Best Local Similarity 96.4%; Score 53; DB 1; Length 502;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQLFNKPYW 9
DB 301 AQLFNKPYW 309

RESULT 15
VL1_HPV26 STANDARD; PRT; 503 AA.
AC P36735;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 26.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
NCBI_TaxID=31549;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H.; Hofmann B.;
RT "Primer-directed sequencing of human Papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74472; CAA52535.1; -.
DR PIR; S36549; S36549.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
DR Coat protein; late_protein.
SQ SEQUENCE 503 AA; 56328 MW; A0706D12F425BE80 CRC64;

Query Match
Best Local Similarity 96.4%; Score 53; DB 1; Length 503;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQLFNKPYW 9
DB 304 AQLFNKPYW 312

RESULT 16
VL1_HPV66 STANDARD; PRT; 503 AA.
AC Q80961;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 66.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.

```

```

OX NCBI_TaxID=37119;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/Genbank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U31794; AAA79505.1; -
DR InterPro; IPR002210; PV_capsid.L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; Late_protein.L1.
DR PRINTS; PR00865; HPVcapsid.L1.
DR ProDom; PD000544; PV_capsid.L1; 1.
DR CoaT protein; Late protein.
DR CoaT protein; Late protein.
SQ SEQUENCE 503 AA; 56174 MW; 20649CE2E7875BA2 CRC64;

Query March 96.4%; Score 53; DB 1; Length 503;
Best Local Similarity 88.9%; Pred. No. 0.014; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 1;

OY 1 AQLFNKPYM 9
Db 303 AQLFNKPYM 311

RESULT 17
ID VLI HPV30 STANDARD; PRT; 508 AA.
AC 002515;
DT 01-JUN-1993 (Rel. 26, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 30.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OC NCBI_TaxID=10611;
OX [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H.; Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
[2]
RP SEQUENCE OF 309-352 FROM N.A.
RX MEDLINE=97407963; PubMed=1326639;
RA Chan S.Y.; Bernard H.U.; Ong C.K.; Chan S.P.; Birgit H.; Delius H.;
RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and
RT variants: a showcase for the molecular evolution of DNA viruses.";
RL J. Virol. 66:5714-5725(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74474; CAAS2548.1; -
DR EMBL; M96279; AAA47018.1; -
DR PIR; S36508; S36508.
DR InterPro; IPR002210; PV_capsid.L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; Late_protein.L1; 1.
OS Human papillomavirus type 30.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

```

```

DR PRINTS; PR00865; HPVcapsid.L1.
DR ProDom; PD000544; PV_capsid.L1; 1.
DR CoaT protein; Late protein.
DR CoaT protein; Late protein.
SQ SEQUENCE 508 AA; 56667 MW; 62F359F257714748 CRC64;

Query March 96.4%; Score 53; DB 1; Length 508;
Best Local Similarity 88.9%; Pred. No. 0.015; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 1;

OY 1 AQLFNKPYM 9
Db 311 AQLFNKPYM 319

RESULT 18
ID VLI HPV10 STANDARD; PRT; 531 AA.
AC P36732;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 10.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OC NCBI_TaxID=10603;
OX [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H.; Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
[2]
CC CAUTION: It is uncertain whether Met-1 or Met-29 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74465; CAAS2494.1; -
DR PIR; S36537; S36537.
DR InterPro; IPR002210; PV_capsid.L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; Late_protein.L1; 1.
DR PRINTS; PR00865; HPVcapsid.L1.
DR ProDom; PD000544; PV_capsid.L1; 1.
DR CoaT protein; Late protein.
DR CoaT protein; Late protein.
SQ SEQUENCE 531 AA; 59011 MW; EDC9BA0DCAY6786 CRC64;

Query March 96.4%; Score 53; DB 1; Length 531;
Best Local Similarity 88.9%; Pred. No. 0.015; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 1;

OY 1 AQLFNKPYM 9
Db 331 AQLFNKPYM 339

RESULT 19
ID VLI HPV56 STANDARD; PRT; 534 AA.
AC P36743;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 56.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

```

```

OC Papillomavirus.
OX NCBI_Taxid=10596;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X74483; CAA52600.1; -.
DR PIR; S36583; S36583.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; Late_protein_L1; 1.
DR PRINTS; PR00865; HPVCAPSIDL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coat protein; late protein.
SQ SEQUENCE 534 AA; 60160 MW; 2ED3A965D1B1EC96 CRC64;

Query Match
Best Local Similarity 96.4%; Score 53; DB 1; Length 534;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPW 9
Db 338 AQIFNKPW 346

RESULT 20
V1 HPV51 STANDARD; PRT; 504 AA.
AC P26536;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 51.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=10595;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=91303675; PubMed=1649326;
RA Lungu O., Crum C.P., Silverstein S.J.;
RT "Biologic properties and nucleotide sequence analysis of human
RT papillomavirus type 51.";
RL J. Virol. 65:4216-4225(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M62877; -. NOT_ANNOTATED_CDS.
DR PIR; G40415; P1ML51.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; Late_protein_L1; 1.
DR PRINTS; PR00865; HPVCAPSIDL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coat protein; late protein.

```

```

SQ SEQUENCE 504 AA; 56314 MW; 4D5940CB31ED9646 CRC64;

Query Match
Best Local Similarity 94.5%; Score 52; DB 1; Length 504;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPW 9
Db 304 AQIFNKPW 312

RESULT 21
V1 HPV54 STANDARD; PRT; 497 AA.
ID V1 HPV54
AC P50819; O81024;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 54.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=37113;
RN (1)
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 310-460 FROM N.A.
RX MEDLINE=95052821; PubMed=7963696;
RA Bernard H.U., Chan S.Y., Manos M.M., Ong C.K., Villa L.L.,
RA Delius H., Peyton C.L., Bauer H.M., Wheeler C.M.;
RT "Identification and assessment of known and novel human
RT papillomaviruses by polymerase chain reaction amplification,
RT restriction fragment length polymorphisms, nucleotide sequence, and
RT phylogenetic algorithms.";
RL J. Infect. Dis. 170:1077-1085(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U37488; AAA79193.1; -.
DR PIR; U12501; AAA67245.1; -.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; Late_protein_L1; 1.
DR PRINTS; PR00865; HPVCAPSIDL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coat protein; late protein.
FT CONFLICT 313 313 Q -> H (IN REF. 2).
FT CONFLICT 323 323 V -> L (IN REF. 2).
FT CONFLICT 368 368 I -> T (IN REF. 2).
FT CONFLICT 375 375 T -> A (IN REF. 2).
FT CONFLICT 439 439 N -> T (IN REF. 2).
FT CONFLICT 456 456 F -> Y (IN REF. 2).
SQ SEQUENCE 497 AA; 55840 MW; A6F7A815F8476A4B CRC64;

Query Match
Best Local Similarity 92.7%; Score 51; DB 1; Length 497;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QIFNKPW 9
Db 239 QIFNKPW 306

RESULT 22

```

```

VL1_HPV33          STANDARD:          PRT:          499 AA.
ID_VL1_HPV33
AC P06416;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 33.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10586;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86200464; PubMed=3009902;
RA Cole S.T., Strecek R.E.;
RT "Genome organization and nucleotide sequence of human papillomavirus
RT type 33, which is associated with cervical cancer.";
RL J. Virol. 58:991-995(1986).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M12732; AAA46964.1; -.
DR PIR; A03641; P1WL33.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsIDL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KM Coat protein; late protein.
SQ SEQUENCE 499 AA; 55902 MW; 28347ADA93DE01D CRC64;

Query Match          90.9%; Score 50; DB 1; Length 499;
Best Local Similarity 77.8%; Pred. No. 0.054;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQTFRKPYW 9
DB 304 SQLEFKPYW 312

RESULT 23
VL1_HPV70          STANDARD:          PRT:          504 AA.
ID_VL1_HPV70
AC P50793;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 70.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=39457;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96249586; PubMed=8915087;
RA Forslund O., Hansson B.G.;
RT "Human papillomavirus type 70 genome cloned from overlapping PCR
RT products: complete nucleotide sequence and genomic organization.";
RL J. Clin. Microbiol. 34:802-809(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97060129; PubMed=8904450;
RA Longuet M., Beaudouin S., Orth G.;
RT "Two novel genital human papillomavirus (HPV) types, HPV68 and HPV70,
RT related to the potentially oncogenic HPV39.";
```

```

RL J. Clin. Microbiol. 34:738-744(1996).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U21941; AAC54857.1; -.
DR EMBL; U22461; AAC54879.1; -.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsIDL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KM Coat protein; late protein.
FT CONFLICT 39 39 S -> T (IN REF. 2).
FT CONFLICT 178 178 T -> TTY (IN REF. 2).
SQ SEQUENCE 504 AA; 56576 MW; 0E9029B5B3F3D6C5 CRC64;

Query Match          90.9%; Score 50; DB 1; Length 504;
Best Local Similarity 77.8%; Pred. No. 0.054;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQTFRKPYW 9
DB 302 SQLEFKPYW 310

RESULT 24
VL1_HPV39          STANDARD:          PRT:          505 AA.
ID_VL1_HPV39
AC P24838;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 39.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91135017; PubMed=1847266;
RA Volpers C., Strecek R.E.;
RT "Genome organization and nucleotide sequence of human papillomavirus
RT type 39.";
RL Virology 181:419-423(1991).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M62849; AAA47056.1; -.
DR PIR; H38502; P1WL39.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsIDL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KM Coat protein; late protein.
SQ SEQUENCE 505 AA; 56604 MW; 73854F9F1C56BCD CRC64;

Query Match          90.9%; Score 50; DB 1; Length 505;
Best Local Similarity 77.8%; Pred. No. 0.054;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy      1 AQTENKPYW 9
Db      303 SOLFNKPYW 311

RESULT 25
ID_VL1_HPV61 STANDARD; PRT; 505 AA.
AC P50822; Q80954;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 61.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=37116;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RT Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RX MEDLINE=95052821; PubMed=7963696;
RA Bernard H.U., Chan S.Y., Manos M.M., Ong C.K., Villa L.L.,
RA Delius H., Peyton C.L., Bauer H.M., Wheeler C.M.;
RT Identification and assessment of known and novel human
RT papillomaviruses by polymerase chain reaction amplification,
RT restriction fragment length polymorphisms, nucleotide sequence, and
RT phylogenetic algorithms";
RL J. Infect. Dis. 170:1077-1085(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U12500; AAA67949.1; -
DR InterPro; IPR002210; PV_Capsid_L1.
DR pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPV_Capsid_L1.
DR ProDom; PD000544; PV_Capsid_L1; 1.
DR KM Coat protein: late protein.
FT CONFLICT 374 374 Q -> P (IN REF. 2).
FT CONFLICT 430 430 Q -> L (IN REF. 2).
SQ SEQUENCE 505 AA; 56710 MW; EB56CA460E94C853 CRC64;

Query Match 90.9%; Score 50; DB 1; Length 505;
Best Local Similarity 77.8%; Pred. No. 0.054;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AQTENKPYW 9
Db      305 SOLFNKPYW 313

RESULT 26
ID_VL1_HPV68 STANDARD; PRT; 505 AA.
AC P54669;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 68.

```

```

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=45240;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97060129; PubMed=8904450;
RA Longuet M., Beaudenon S., Orth G.;
RT "Two novel genital human papillomavirus (HPV) types, HPV68 and HPV70,
RT related to the potentially oncogenic HPV39.";
RL J. Clin. Microbiol. 34:738-744(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X67161; CAA47634.1; -
DR InterPro; IPR002210; PV_Capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPV_Capsid_L1.
DR ProDom; PD000544; PV_Capsid_L1; 1.
DR KM Coat protein: late protein.
SQ SEQUENCE 505 AA; 56693 MW; 246CDB4DBA2A1F97 CRC64;

Query Match 90.9%; Score 50; DB 1; Length 505;
Best Local Similarity 77.8%; Pred. No. 0.054;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AQTENKPYW 9
Db      304 SOLFNKPYW 312

RESULT 27
ID_VL1_HPVME STANDARD; PRT; 505 AA.
AC P27964;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type ME180.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=10602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9137416; PubMed=1716694;
RA Reuter S., Delius H., Kahn T., Hofmann B., Zur Hausen H., Schwarz E.;
RT "Characterization of a novel human papillomavirus DNA in the cervical
RT carcinoma cell line ME180.";
RL J. Virol. 65:5564-5568(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M73258; AAF14010.1; -
DR PIR; BA0509; P1W1BR.
DR InterPro; IPR002210; PV_Capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPV_Capsid_L1.
DR ProDom; PD000544; PV_Capsid_L1; 1.

```


KW Coat protein; Late protein.
SQ SEQUENCE 505 AA; 56805 MW; 595DE493B708207B CRC64;

Query Match 90.9%; Score 50; DB 1; Length 505;
Best Local Similarity 77.8%; Pred. No. 0.054;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 A0IENKPYW 9
DB 304 SQLEFKPYW 312

RESULT 28

VL1_HPV58 STANDARD; PRT; 524 AA.
ID_VL1_HPV58
AC P26535;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 58.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024102; PubMed=1656594;
RA Kiri Y., Iwamoto S., Matsukura T.;
RT "Human papillomavirus type 58 DNA sequence.";
RL Virology 185:424-427(1991).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL; D90400; BAA31851.1; -.
DR PIR; G36779; P1WL58.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; Late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coat protein; Late protein.
SQ SEQUENCE 524 AA; 59038 MW; 5D8C95E2878A26DF CRC64;

Query Match 90.9%; Score 50; DB 1; Length 524;
Best Local Similarity 77.8%; Pred. No. 0.056;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 A0IENKPYW 9
DB 330 SQLEFKPYW 338

RESULT 29
VL1_HPV52 STANDARD; PRT; 529 AA.
ID_VL1_HPV52
AC Q05138;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 52.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10618;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).

RN [2]
RP SEQUENCE OF 333-376 FROM N.A.
RX MEDLINE=92407963; PubMed=1326639;
RA Chan S.Y., Bernard H.U., Ong C.K., Chan S.P., Birgit H., Delius H.;
RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and
RT variants: a showcase for the molecular evolution of DNA viruses.";
RL J. Virol. 66:5714-5725(1992).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL; X74481; CAAS2590.1; -.
DR EMBL; M96297; AAA47036.1; -.
DR PIR; S36578; S36578.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; Late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coat protein; Late protein.
SQ SEQUENCE 529 AA; 59469 MW; 77F7DB5CA9664C53 CRC64;

Query Match 90.9%; Score 50; DB 1; Length 529;
Best Local Similarity 77.8%; Pred. No. 0.057;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 A0IENKPYW 9
DB 335 SQLEFKPYW 343

RESULT 30
VL1_HPV45 STANDARD; PRT; 539 AA.
ID_VL1_HPV45
AC P36741;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 45.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL; X74479; CAAS2578.1; -.
DR PIR; S36566; S36566.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 KM Coat protein; Late protein.
 SQ SEQUENCE 539 AA; 60310 MW; 9D79FF68042345E CRC64;

Query Match 90.9%; Score 50; DB 1; Length 539;
 Best Local Similarity 77.8%; Pred. No. 0.058;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 AQIFNKPW 9
 Db 333 SQIFNKPW 341

RESULT 31
 ID VLI_HPV18 STANDARD; PRT; 568 AA.
 AC P06794; O84270;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Major capsid protein L1.
 GN L1.
 OS Human papillomavirus type 18.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10582;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87283882; PubMed=3039146;
 RA Cole S.T., Danos O.;
 RT "Nucleotide sequence and comparative analysis of the human
 papillomavirus type 18 genome. Phylogeny of papillomaviruses and
 repeated structure of the E6 and E7 gene products.";
 RT J. Mol. Biol. 193:599-608(1987).
 RL [2]
 RP SEQUENCE OF 363-406 FROM N.A.
 RX MEDLINE=92407963; PubMed=1326639;
 RA Chan S.Y., Bernard H.V., Ong C.K., Chan S.P., Birgit H., Delius H.;
 RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and
 variants: a showcase for the molecular evolution of DNA viruses.";
 RT J. Virol. 66:5714-5725(1992).
 RL [3]
 RP SEQUENCE OF 501-568 FROM N.A.
 RX MEDLINE=87061231; PubMed=3023691;
 RA Thierry F., Heard J.M., Dartmann K., Yaniv M.;
 RT "Characterization of a transcriptional promoter of human
 papillomavirus 18 and modulation of its expression by simian virus 40
 and adenovirus early antigens.";
 RT J. Virol. 61:134-142(1987).
 RL -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X05015; CA28671.1; -;
 DR EMBL; M96287; AAA47026.1; -;
 DR EMBL; M14710; AAA65508.1; -;
 DR EMBL; A06329; CAA00545.1; -;
 DR PIR; A26251; P1ML18.
 DR InterPro; IPR002210; PV_capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 DR Coat protein; Late protein.
 FT CONFLICT 384 384 V -> I (IN REF. 2).
 SO SEQUENCE 568 AA; 63623 MW; BB2B6F61177FC10 CRC64;

Query Match 90.9%; Score 50; DB 1; Length 568;
 Best Local Similarity 77.8%; Pred. No. 0.061;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 AQIFNKPW 9
 Db 365 SQIFNKPW 373

RESULT 32
 ID VLI_HPV2A STANDARD; PRT; 510 AA.
 AC P25486;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Major capsid protein L1.
 GN L1.
 OS Human papillomavirus type 2a.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10584;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91188699; PubMed=1964523;
 RA Hirsch-Behnam A., Delius H., de Villiers E.M.;
 RT "A comparative sequence analysis of two human papillomavirus (HPV)
 types 2a and 57.";
 RT Virus Res. 18:81-98(1990).
 RL -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X55964; -; NOT_ANNOTATED_CDS.
 DR PIR; S15620; S15620.
 DR InterPro; IPR002210; PV_capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 DR Coat protein; Late protein.
 SQ SEQUENCE 510 AA; 57193 MW; 83509AFB9DAF70D9 CRC64;
 Query Match 89.1%; Score 49; DB 1; Length 510;
 Best Local Similarity 87.5%; Pred. No. 0.085;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

RX MEDLINE=91188699; PubMed=1964523;
RA Hirsch-Benham A., Delius H., de Villiers E.M.;
RT "A comparative sequence analysis of two human papillomavirus (HPV)
RL types 2a and 57."
CC -----
CC Virus Res. 18:81-98(1990).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X55965; CAA39436.1; -.
CC PIR; S15627; S15627.
CC InterPro: IPR002210; PV_capsid_L1.
CC InterPro: IPR008975; Viral_cap_coat.
CC Pfam; PF00500; late_protein_L1; 1.
CC PRINTS; PR00865; HPV_CapsID1.
CC ProDom; PD000544; PV_capsid_L1; 1.
CC Coats protein; Late protein.
CC SEQUENCE 510 AA; 57239 MW; C35E9B0A5AB2B6F5 CRC64;
SQ
Query Match 89.1%; Score 49; DB 1; Length 510;
Best Local Similarity 87.5%; Pred. No. 0.085; 0; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 0;
OY 2 QIFNKPYW 9
ID_VL1_HPV03 STANDARD; PRT; 532 AA.
AC P36731.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 3.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OC NCBI_TaxID=10614;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -1- CAUTION: It is uncertain whether Met-1 or Met-29 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X74462; CAA52474.1; -.
CC PIR; S36554; S36554.
CC InterPro: IPR002210; PV_capsid_L1.
CC InterPro: IPR008975; Viral_cap_coat.
CC Pfam; PF00500; late_protein_L1; 1.
CC PRINTS; PR00865; HPV_CapsID1.
CC ProDom; PD000544; PV_capsid_L1; 1.
CC Coats protein; Late protein.
CC SEQUENCE 532 AA; 59194 MW; A0AFB87DBEE98B27 CRC64;
SQ
Query Match 89.1%; Score 49; DB 1; Length 532;

```

```

Best Local Similarity 87.5%; Pred. No. 0.089; 0; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 0;
OY 2 QIFNKPYW 9
ID_VL1_HPV27 STANDARD; PRT; 594 AA.
AC P36736; Q81959;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 27.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OC NCBI_TaxID=31550;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-110 IS THE
CC INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X74473; CAA52541.1; -.
CC PIR; S36502; S36502.
CC InterPro: IPR002210; PV_capsid_L1.
CC InterPro: IPR008975; Viral_cap_coat.
CC Pfam; PF00500; late_protein_L1; 1.
CC PRINTS; PR00865; HPV_CapsID1.
CC ProDom; PD000544; PV_capsid_L1; 1.
CC Coats protein; Late protein.
CC SEQUENCE 594 AA; 67098 MW; 946BA35505DDC0A8 CRC64;
SQ
Query Match 89.1%; Score 49; DB 1; Length 594;
Best Local Similarity 87.5%; Pred. No. 0.1; 0; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 0;
OY 2 QIFNKPYW 9
ID_VL1_BPVL STANDARD; PRT; 495 AA.
AC P03103;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Bovine papillomavirus type 1.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OC NCBI_TaxID=10559;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=83012974; PubMed=6289124;
RA Chen E.Y., Howley P.M., Levinson A.D., Seeburg P.H.;
RT "The primary structure and genetic organization of the bovine
RL papillomavirus type 1 genome."
CC Nature 299:529-534(1982).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X02346; CAB46515.1; -
DR EMBL; A14693; CAA01196.1; -
DR PIR; A03644; P1MLB.
DR InterPro; IPR002210; PV_capsid_L1.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
DR Coats protein; Late protein.
KW SEQUENCE 495 AA; 55551 MW; EA4B379D1CD1C51 CRC64;
SQ
Query Match
Best Local Similarity 87.3%; Score 48; DB 1; Length 495;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 QIFNRPYW 9
DB 302 QIFNRPYW 309
-----
RESULT 37
ID -VLI_BPV2 STANDARD; PRT; 497 AA.
AC P06458;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Bovine papillomavirus type 2.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OC NCBI_TaxID=10560;
OK [1]
RA SEQUENCE FROM N.A.
RA Groff D.E., Mitera R., Lancaster W.D.;
RL Submitted (MAY-1988) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85107102; PubMed=2981958;
RA Potter H.L., Weinke W.J.;
RT "Nucleotide sequence of bovine papillomavirus type 2 late region."
RL J. Gen. Virol. 66:187-193(1985).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M20219; AAA66840.1; -
DR EMBL; X01768; CAA25908.1; -
DR PIR; A31169; P1MLB2.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.

```

```

DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coats protein; Late protein.
FT CONFLICT 176 176 L -> I (IN REF. 2).
FT CONFLICT 323 323 N -> D (IN REF. 2).
FT CONFLICT 386 386 L -> V (IN REF. 2).
FT CONFLICT 442 442 K -> R (IN REF. 2).
FT SEQUENCE 497 AA; 55595 MW; 674728E3A884A068 CRC64;
SQ
Query Match
Best Local Similarity 87.3%; Score 48; DB 1; Length 497;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 QIFNRPYW 9
DB 302 QIFNRPYW 309
-----
RESULT 38
ID -VLI_HPV63 STANDARD; PRT; 507 AA.
AC Q07861;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 63.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OC NCBI_TaxID=28311;
OK [1]
RA SEQUENCE FROM N.A.
RX MEDLINE=93276568; PubMed=8389082;
RA Egawa K., Delius H., Matsukura T., Kawashima M., de Villiers E.M.;
RT "Two novel types of human papillomavirus, HPV 63 and HPV 65:
RT comparisons to other HPV types."
RL Virology 194:789-799(1993).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X70828; CAA50170.1; -
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
DR Coats protein; Late protein.
KW SEQUENCE 507 AA; 57555 MW; 4466333C9C9403A9 CRC64;
SQ
Query Match
Best Local Similarity 87.3%; Score 48; DB 1; Length 507;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 QIFNRPYW 9
DB 310 QIFNRPYW 317
-----
RESULT 39
ID -VLI_HPV48 STANDARD; PRT; 513 AA.
AC P50817; O80926;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major capsid protein L1.

```

```

GN L1.
OS Human papillomavirus type 48.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=40538;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 377-471 FROM N.A.
RA Chan S.Y., Delius H., Halpern A.L., Bernard H.U.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U31789; AAA79470.1; -.
DR EMBL; U21869; AAA92831.1; -.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KM Coat protein; Late protein.
SQ SEQUENCE 513 AA; 58012 MW; 09C78670BEDC92CE CRC64;

Query Match 85.5%; Score 47; DB 1; Length 513;
Best Local Similarity 66.7%; Pred. No. 0.21;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQLFNKPYW 9
Db 309 SQLFNRPYW 317

RESULT 40
VLI_HPV50
ID VLI_HPV50 STANDARD; PRT; 515 AA.
AC P50818; Q80933;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 50.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=40539;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 380-475 FROM N.A.
RA Chan S.Y., Delius H., Halpern A.L., Bernard H.U.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U31790; AAA79477.1; -.
DR EMBL; U21870; AAA92832.1; -.

```

```

DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KM Coat protein; Late protein.
SQ SEQUENCE 515 AA; 58425 MW; 0D79EA3A1A215BC0 CRC64;

Query Match 85.5%; Score 47; DB 1; Length 515;
Best Local Similarity 66.7%; Pred. No. 0.21;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQLFNKPYW 9
Db 312 SQLFNRPYW 320

Search completed: March 3, 2004, 10:23:19
Job time : 11 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: March 3, 2004, 10:21:39 ; Search time 13.5 Seconds
(without alignments)
64.128 Million cell updates/sec

Title: US-09-980-064-2
Perfect score: 49
Sequence: 1 AGVDNRECT 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	49	100.0	531	1 P1WLS	major capsid prote
2	42	85.7	502	2 S36526	L1 protein - human
3	42	85.7	505	1 P1WL35	L1 protein - human
4	42	85.7	529	2 S36578	L1 protein - human
5	41	83.7	504	1 P1WL31	L1 protein - human
6	39	79.6	499	1 P1WL33	L1 protein - human
7	37	75.5	418	2 G69302	sulfite reductase
8	37	75.5	524	1 P1WL58	L1 protein - human
9	36	73.5	474	2 T23547	hypothetical prote
10	36	73.5	1039	2 G66203	probable N-arginin
11	35	71.4	243	2 AB0733	probable bacteriop
12	35	71.4	1737	2 AS9235	unconventional myo
13	34	69.4	247	2 AC2699	conserved hypotet
14	34	69.4	260	2 C97481	1828 [imported] -
15	34	69.4	596	2 T44348	GRP binding protei
16	34	69.4	831	2 T25827	hypothetical prote
17	34	69.4	2471	2 A49128	cell-fate determin
18	34	69.4	2531	2 S18188	notch protein homo
19	34	69.4	2551	2 A46019	notch-1 protein -
20	34	69.4	2555	2 A46043	notch protein -
21	33	67.3	233	1 K1MST	thymidine kinase (
22	33	67.3	307	2 T27332	hypothetical prote
23	33	67.3	322	2 T27333	hypothetical prote
24	33	67.3	575	2 F83904	hypothetical prote
25	33	67.3	688	2 T33708	hypothetical prote
26	33	67.3	870	2 G86450	F5D14_31 protein -
27	33	67.3	984	1 A34076	protein-cytosine k
28	33	67.3	1101	2 S51823	myosin heavy chain
29	33	67.3	1652	2 T16799	hypothetical prote

30	32	65.3	59	1 FEMZB	ferredoxin 2[4Fe-4
31	32	65.3	91	2 F83636	hypothetical prote
32	32	65.3	134	2 S69360	retinol-binding pr
33	32	65.3	135	1 RJHDO	retinol-binding pr
34	32	65.3	135	1 RJRTO	retinol-binding pr
35	32	65.3	135	2 S16355	retinol-binding pr
36	32	65.3	213	2 B66001	probable DNA-direc
37	32	65.3	243	2 AG3251	conserved hypotet
38	32	65.3	264	1 A40019	osteopontin precu
39	32	65.3	264	2 F83351	conserved hypotet
40	32	65.3	268	2 E83020	lipopolysaccharide
41	32	65.3	308	2 T44427	hypothetical prote
42	32	65.3	308	2 T43293	multidrug resistan
43	32	65.3	393	2 C97213	electron-transferr
44	32	65.3	442	2 S20051	hypothetical prote
45	32	65.3	457	2 F70885	probable glna4 pro

ALIGNMENTS

RESULT 1
P1WLS
major capsid protein L1 - human papillomavirus type 16
C:Species: human papillomavirus type 16
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 20-Aug-1999
C:Accession: A03640; F44906; T10431
R:Seedorf, K.; Krammer, G.; Durst, M.; Suhai, S.; Roweckamp, W.G.
Virology 145, 181-185, 1985
A>Title: Human papillomavirus type 16 DNA sequence.
A:Reference number: A22355; MUID:85246220; PMID:2990099
A:Accession: A03640
A:Molecule type: DNA
A:Residues: 1-531 <SEB>
A:Cross-references: GB:K02718; NID:G333031; PIDN:AAA46943.1; PID:G333037
R:van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnenackers, H.F.; Delius, H
J. Clin. Microbiol. 30, 1716-1721, 1992
A>Title: General primer polymerase chain reaction in combination with sequence analysis
A:Reference number: A44889; MUID:92332706; PMID:1321168
A:Contents: mucosotropic type 16, cervical smear
A:Accession: F44906
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 365-395 <VAN>
A>Note: sequence extracted from NCBI backbone (NCBIP:109407)
R:Kennedy, I.M.; Haddow, J.K.; Clements, J.B.
J. Virol. 65, 2093-2097, 1991
A>Title: A negative element in the human poapillomavirus type 16 genome acts at the leve
A:Reference number: 217014; MUID:91162763; PMID:1848319
A:Accession: T10431
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-531 <KEN>
A:Cross-references: EMBL:K02718; NID:G333031; PIDN:AAA46943.1; PID:G333037
C:Genetics:
A:Gene: L1
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein
Query Match 100.0%; Score 49; DB 1; Length 531;
Best local similarity 100.0%; Pred. No. 0.071;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGVDNRECT 9
DB 165 AGVDNRECT 173
RESULT 2
S36526
L1 protein - human papillomavirus type 35H
C:Species: human papillomavirus type 35H
C:Date: 09-Dec-1993 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999

C:Accession: S36526
R:Delius, H.; Hofmann, B.
Submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36526
A:Molecule type: DNA
A:Residues: 1-502
A:Cross-references: EMBL:X74477; NID:g396997; PIDN:CAA52566.1; PID:g397004
A:Experimental source: strain 35H
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match 85.7%; Score 42; DB 2; Length 502;
Best Local Similarity 77.8%; Pred. No. 1.7;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 AGVDNRECI 9
DB 139 SGTDRNRECI 147

RESULT 3

PIWL35
L1 protein - human papillomavirus type 35
C:Species: human papillomavirus type 35
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jul-1999
C:Accession: G40824; C44889
R:Marich, J.B.; Pontsler, A.V.; Rice, S.M.; McGraw, K.A.; Dubensky, T.W.
Virology 186, 770-776, 1992
A:Title: The phylogenetic relationship and complete nucleotide sequence of human papillomavirus type 35
A:Reference number: A40824; MUID:92124753; PMID:1310198
A:Accession: G40824
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-505 <VAR>
A:Cross-references: GB:M74117; NID:g333050; PIDN:AAA6972.1; PID:g333058
R:van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnemakers, H.F.; Delius, H.
J. Clin. Microbiol. 30, 1716-1721, 1992
A:Title: General primer polymerase chain reaction in combination with sequence analysis
A:Reference number: A44889; MUID:92332706; PMID:1321168
A:Accession: C44889
A:Molecule type: DNA
A:Residues: 338-370 <VAN>
A:Cross-references: GB:S40240; NID:g251686; PIDN:AA22564.1; PID:g251687
A:Experimental source: mucosotropic type 35, cervical smear
A:Note: sequence extracted from NCBI backbone (NCBIN:109393, NCBIP:109408)
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match 85.7%; Score 42; DB 1; Length 505;
Best Local Similarity 77.8%; Pred. No. 1.7;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 AGVDNRECI 9
DB 142 SGTDRNRECI 150

RESULT 4

S36578
L1 protein - human papillomavirus type 52
C:Species: human papillomavirus type 52
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36578; I44889
R:Delius, H.; Hofmann, B.
Submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36578
A:Molecule type: DNA
A:Residues: 1-529

A:Cross-references: EMBL:X74481; NID:g397038; PIDN:CAA52590.1; PID:g397045
R:van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnemakers, H.F.; Delius, H.
J. Clin. Microbiol. 30, 1716-1721, 1992
A:Title: General primer polymerase chain reaction in combination with sequence analysis
A:Reference number: A44889; MUID:92332706; PMID:1321168
A:Accession: I44889
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 368-399 <VAN>
A:Cross-references: GB:S40277; NID:g251698; PIDN:AA22570.1; PID:g251699
A:Experimental source: type Xa, cervical carcinoma in situ
A:Note: sequence extracted from NCBI backbone (NCBIN:109399, NCBIP:109447)
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match 85.7%; Score 42; DB 2; Length 529;
Best Local Similarity 75.0%; Pred. No. 1.8;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 GVDNRECI 9
DB 169 GTDRNRECI 176

RESULT 5

PIWL31
L1 protein - human papillomavirus type 31
C:Species: human papillomavirus type 31
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C:Accession: G32444
R:Goldborough, M.D.; Disilverstre, D.; Temple, G.F.; Lorincz, A.T.
Virology 171, 306-311, 1989
A:Title: Nucleotide sequence of human papillomavirus type 31: a cervical neoplasia-assoc
A:Reference number: A94398; MUID:89299478; PMID:2545036
A:Accession: G32444
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-504 <GOL>
A:Cross-references: GB:J04353; NID:g333048; PIDN:AAA46956.1; PID:g459922
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match 83.7%; Score 41; DB 1; Length 504;
Best Local Similarity 87.5%; Pred. No. 2.7;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GVDNRECI 9
DB 141 GTDRNRECI 148

RESULT 6

PIWL33
L1 protein - human papillomavirus type 33
C:Species: human papillomavirus type 33
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03641
R:Cole, S.T.; Strecek, R.E.
J. Virol. 58, 991-995, 1986
A:Title: Genome organization and nucleotide sequence of human papillomavirus type 33, wh
A:Reference number: A93020; MUID:86200464; PMID:3009302
A:Accession: A03641
A:Molecule type: DNA
A:Residues: 1-499 <COL>
A:Cross-references: GB:M12732; NID:g333049; PIDN:AAA46964.1; PID:g463183
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match 79.6%; Score 39; DB 1; Length 499;
Best Local Similarity 75.0%; Pred. No. 6.7;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GVDNRECI 9
 Db 141 GADNRECI 148

RESULT 7

669302
 sulfite reductase (EC 1.8.99.1) alpha chain dera - Archaeoglobus fulgidus
 C/Species: Archaeoglobus fulgidus
 C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
 C/Accession: G69302; 139661; S27478
 R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirschner, E.F.
 .; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 .; Nature 390, 364-370, 1997
 A:Authors: Uterback, T.; Colton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
 .; Smith, H.O.; Weese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: G69302
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-418 <KGE>
 A:Cross-references: GB:AB001075; GB:AB000782; NID:G2689398; PIDN:AB90812.1; PID:G265020
 R/Dahl, C.; Kredich, N.M.; Deutmann, R.; Truuper, H.G.
 J. Gen. Microbiol. 139, 1817-1828, 1993
 A:Title: Dissimilatory sulphite reductase from Archaeoglobus fulgidus: physicochemical
 A:Reference number: I39661; MUID:94014979; PMID:7691984
 A:Accession: I39661
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-418 <RES>
 A:Cross-references: GB:M95624; NID:G142188; PIDN:AB17213.1; PID:G142189
 C/Genetics:
 A:Gene: dsrA
 C:Keywords: oxidoreductase

Query Match 75.5%; Score 37; DB 2; Length 418;
 Best Local Similarity 71.4%; Pred. No. 14;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 3 VDNRECI 9
 Db 281 IDNRECI 287

RESULT 8

PIWL58
 LI protein - human papillomavirus type 58
 C/Species: human papillomavirus type 58
 A/Note: host Homo sapiens (man)
 C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jul-2000
 C/Accession: G36779
 R/Kirli, Y.; Iwamoto, S.; Matsukura, T.
 Virolgy 185, 424-427, 1991
 A:Title: Human papillomavirus type 58 DNA sequence.
 A:Reference number: A36779; MUID:92024102; PMID:1656594
 A:Accession: G36779
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-524 <KIR>
 A:Cross-references: GB:D90400; NID:G222386; PIDN:BAA31851.1; PID:G3337104
 C/Superfamily: Papillomavirus LI protein
 C:Keywords: late protein

Query Match 75.5%; Score 37; DB 1; Length 524;
 Best Local Similarity 75.0%; Pred. No. 18;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GVDNRECI 9
 Db 167 GSDNRECI 174

RESULT 9

123547
 hypothetical protein K09E9.3 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C/Accession: T23547
 R/Swinburne, J.; Ainscough, R.
 submitted to the EMBL Data Library, August 1996
 A:Reference number: 219757
 A:Accession: T23547
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-474 <WIL>
 A:Cross-references: EMBL:Z79602; PIDN:CA01890.1; GSPDB:GN00028; CESP:K09E9.3
 A:Experimental source: clone K09E9
 C/Genetics:
 A:Gene: CESP:K09E9.3
 A:Map position: X
 A:Introns: 23/3; 99/3; 172/1; 206/2; 445/1

Query Match 73.5%; Score 36; DB 2; Length 474;
 Best Local Similarity 66.7%; Pred. No. 25;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 AGVDNRECI 9
 Db 275 AGIDNRECI 283

RESULT 10

686203
 Probable N-arginine diase convertase [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Oct-2003
 C/Accession: G86203
 R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Matti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: G86203
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1039 <STO>
 A:Cross-references: GB:AB005172; NID:G7523693; PIDN:ABF63132.1; GSPDB:GN00141
 C/Genetics:
 A:Map position: 1
 C:Superfamily: insulin-degrading enzyme (IDE)

Query Match 73.5%; Score 36; DB 2; Length 1039;
 Best Local Similarity 87.5%; Pred. No. 54;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GVDNRECI 9
 Db 271 GVDNRECI 278

RESULT 11

AB0733
 Probable bacteriophage protein STY2021 [imported] - Salmonella enterica subsp. enterica
 C/Species: Salmonella enterica subsp. enterica serovar Typhi
 A/Note: this species has also been called Salmonella typhi
 C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C/Accession: AB0733
 R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
 A:Reference number: AB0502; MWID:21534947; PMID:11677608
 A:Accession: AB0733
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-243 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD05564.1; PID:gl6503061; GSPDB:GN00176
 C:Genetics:
 A:Gene: SRY2021

Query Match 71.4%; Score 35; DB 2; Length 243;
 Best Local Similarity 55.6%; Pred. No. 21;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
 ||:|||||
 DB 227 AGLENSECV 235

RESULT 12
 A59235
 unconventional myosin heavy chain Myom - slime mold (*Dictyostelium discoideum*)
 C:Species: *Dictyostelium discoideum*
 C:Date: 19-May-2000 #sequence_rev15 19-May-2000 #text_change 08-Sep-2000
 C:Accession: A59235
 R:Geisler, H.; Schwarz, E.C.; Soldati, T.
 submitted to Genbank, September 1998
 A:Description: Identification of two novel and highly divergent myosins in *Dictyostelium*
 A:Reference number: A59235
 A:Accession: A59235
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-1173 <GR>
 A:Cross-references: GB:AF090533; NID:gs714395; PIDN:ADD47903.1; PID:gs714396
 A:Experimental source: strain AX2
 C:Genetics:
 A:Gene: myom
 A:Map position: 6, aldb-cabA2
 C:Superfamily: myosin motor domain homology #status atypical <MNO>
 F:62-674/Domain: myosin motor domain homology

Query Match 71.4%; Score 35; DB 2; Length 1737;
 Best Local Similarity 71.4%; Pred. No. 1,4e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNRECI 9
 ||:|||||
 DB 531 IDNKECI 537

RESULT 13
 AC2699
 conserved hypothetical protein Atu0995 [imported] - *Agrobacterium tumefaciens* (strain C58)
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 11-Jan-2002 #sequence_rev15 11-Jan-2002 #text_change 18-Nov-2002
 C:Accession: AC2699
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan,
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A:Reference number: AB2577; MWID:21608550; PMID:11743193
 A:Accession: AC2699
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-247 <KUR>
 A:Cross-references: GB:AE008668; PIDN:AAL4209.1; PID:gl7739383; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:
 A:Gene: Atu0995
 A:Map position: circular chromosome

Query Match 69.4%; Score 34; DB 2; Length 247;
 Best Local Similarity 75.0%; Pred. No. 34;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGVDNREC 8
 ||:|||||
 DB 46 AGIDWREC 53

RESULT 14
 C97481
 1828 [imported] - *Agrobacterium tumefaciens* (strain C58, Cereon)
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 30-Sep-2001 #sequence_rev15 30-Sep-2001 #text_change 18-Nov-2002
 C:Accession: C97481
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.,
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
 A:Reference number: A97359; MWID:21608551; PMID:11743194
 A:Accession: C97481
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-260 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK6804.1; PID:gs156010; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C_1828
 A:Map position: circular chromosome

Query Match 69.4%; Score 34; DB 2; Length 260;
 Best Local Similarity 75.0%; Pred. No. 36;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGVDNREC 8
 ||:|||||
 DB 59 AGIDWREC 66

RESULT 15
 T44348
 GRP binding protein [imported] - *Clostridium histolyticum*
 C:Species: *Clostridium histolyticum*
 C:Date: 21-Jan-2000 #sequence_rev15 21-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T44348
 R:Matsumita, O.; Jung, C.M.; Katayama, S.; Minami, J.; Takahashi, Y.; Okabe, A.
 J. Bacteriol. 181, 923-933, 1999
 A:Title: Gene duplication and multiplicity of collagenases in *Clostridium histolyticum*.
 A:Reference number: Z22752; MWID:99121032; PMID:9922257
 A:Accession: T44348
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-536 <MAT>
 A:Cross-references: EMBL:AB014075; NID:gs3868863; PIDN:BAA34254.1; PID:gs3868866
 A:Experimental source: strain JCM 1403
 C:Genetics:
 A:Gene: hflX

Query Match 69.4%; Score 34; DB 2; Length 596;
 Best Local Similarity 75.0%; Pred. No. 80;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVDNRECI 9
 ||:|||||
 DB 482 GVDPRKECI 489

RESULT 16
 T25827

hypothetical protein K12D9.12 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T58827
R:Graves, T.
submitted to the EMBL Data Library, November 1996
A:Description: The sequence of C. elegans coemid K12D9.
A:Reference number: 220093
A:Accession: T58827
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-831 <GRA>
A:Cross-references: EMBL:U80030; P1DN:AA837610.1; GSPDB:GN00023; CESP:K12D9.12
C:Genetics:
A:Gene: CESP:K12D9.12
A:Map position: 5
A:Introns: 38/1; 122/1; 188/1; 209/1; 356/2; 374/1; 458/1; 524/1; 621/3; 746/1; 776/2; 7

Query Match 69.4%; Score 34; DB 2; Length 831;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DNRECT 9
Db 149 IDNRQCV 154

RESULT 17
A49128
cell-fate determining gene Notch2 protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 02-Aug-2002
C:Accession: A49128
R:Weinmaster, G.; Roberts, V.J.; Lemke, G.
Development 116; 931-941, 1992
A:Title: Notch2: a second mammalian Notch gene.
A:Reference number: A49128; MUID:93202015; PMID:1295745
A:Accession: A49128
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-2471 <WEI>
A:Experimental source: Schwann cell
A>Note: Sequence extracted from NCBI backbone (NCBIP:127811)
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology
F:264-295/Domain: EGF homology <EGF1>
F:799-830/Domain: EGF homology <EGF1>
F:877-908/Domain: EGF homology <EGF2>
F:1029-1060/Domain: EGF homology <EGF>
F:1067-1098/Domain: EGF homology <EGF3>
F:1153-1184/Domain: EGF homology <EGF3>
F:1191-1222/Domain: EGF homology <EGF4>
F:1876-1908/Domain: ankyrin repeat homology <AN1>
F:1909-1941/Domain: ankyrin repeat homology <AN2>
F:1943-1975/Domain: ankyrin repeat homology <AN3>
F:1976-2008/Domain: ankyrin repeat homology <AN4>
F:2009-2041/Domain: ankyrin repeat homology <AN5>

Query Match 69.4%; Score 34; DB 2; Length 2471;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VDNRCT 9
Db 1627 IDNRQCV 1633

RESULT 18
S18188
notch protein homolog - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 02-Aug-2002
C:Accession: S18188

R:Weinmaster, G.; Roberts, V.J.; Lemke, G.
Development 113; 199-205, 1991
A:Title: A homolog of Drosophila Notch expressed during mammalian development.
A:Reference number: S18188; MUID:92111383; PMID:1764995
A:Accession: S18188
A:Molecule type: mRNA
A:Residues: 1-2531 <WEI>
A:Cross-references: EMBL:X57405; NID:957634; PID:957635
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology
F:987-1018/Domain: EGF homology <EGF1>
F:1025-1056/Domain: EGF homology <EGF2>
F:1233-1264/Domain: EGF homology <EGF2>
F:1917-1949/Domain: ankyrin repeat homology <AN1>
F:1950-1982/Domain: ankyrin repeat homology <AN2>
F:1984-2016/Domain: ankyrin repeat homology <AN3>
F:2017-2049/Domain: ankyrin repeat homology <AN4>
F:2050-2082/Domain: ankyrin repeat homology <AN5>

Query Match 69.4%; Score 34; DB 2; Length 2531;
Best Local Similarity 57.1%; Pred. No. 3.2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VDNRCT 9
Db 1670 IDNRQCV 1676

RESULT 19
A46019
notch-1 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 07-Mar-2003
C:Accession: A46019; S25144; C9175; B4638; A46438; PH1569; S32109
R:del Amo, F.F.; Gendron-Maguire, M.; Swiatek, P.J.; Jenkins, N.A.; Copeland, N.G.; Grid
Genomics 15; 259-264, 1993
A:Title: Cloning, analysis, and chromosomal localization of Notch-1, a mouse homolog of
A:Reference number: A46019; MUID:93194170; PMID:8449489
A:Accession: A46019
A:Status: not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-2531
A:Cross-references: GB:Z11886; GB:S47228; NID:9288502; P1DN:CAA77941.1; PID:9288503
A>Note: Sequence extracted from NCBI backbone (NCBIP:127318)
R:Ranco del Amo, F.; Smith, D.E.; Swiatek, P.J.; Gendron-Maguire, M.; Greenspan, R.J.; ;
submitted to the EMBL Data Library, April 1992
A:Description: Expression pattern of Notch, a mouse homolog of Drosophila Notch, suggest
A:Reference number: S25144
A:Accession: S25144
A:Molecule type: mRNA
A:Residues: 1551-2108, 'Q', 2110-2114, 'ALP', 2118-2170 <FRA>
A:Cross-references: EMBL:X11886
R:Lardeelli, M.; Lendahl, U.
Exp. Cell Res. 204; 364-372, 1993
A:Title: Notch A and Notch B--two mouse Notch homologues coexpressed in a wide variety o
A:Reference number: A49175; MUID:93178563; PMID:8440332
A:Accession: C49175
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1161-1547 <LAR>
A:Cross-references: EMBL:X68278; NID:9287987; P1DN:CAA48339.1; PID:9287988
A:Experimental source: embryo
A>Note: Sequence extracted from NCBI backbone (NCBIP:126159)
R:Kopan, R.; Weintraub, H.
J. Cell Biol. 121; 631-641, 1993
A:Title: Mouse notch: expression in hair follicles correlates with cell fate determinati
A:Reference number: A46438; MUID:93252998; PMID:8486742
A:Accession: B46438
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1865-1932, 'RR', 1935-1937, 'L', 1938-1967, 'T', 1969-2044, 'IE', 2047-2052, 'S', 2054
A:Experimental source: embryo
A>Note: Sequence extracted from NCBI backbone (NCBIP:131246, NCBIP:131247)

C:Comment: This protein has many EGF repeats and lin-12[1172]/Notch repeats.
 C:Comment: This protein is one of the neurogenic proteins controlling the decision between
 A:Gene: notch-1
 A:Map position: 2
 A:Note: proximal region of chromosome 2
 C:Superfamily: notch protein; ankyrin repeat homology; EGF homology
 F:106-138/Domain: EGF homology <EGF1>
 F:144-175/Domain: EGF homology <EGF2>
 F:222-254/Domain: EGF homology <EGF2>
 F:261-292/Domain: EGF homology <EGF2>
 F:339-370/Domain: EGF homology <EGF3>
 F:416-449/Domain: EGF homology <EGF3>
 F:456-487/Domain: EGF homology <EGF4>
 F:494-525/Domain: EGF homology <EGF5>
 F:532-563/Domain: EGF homology <EGF6>
 F:607-638/Domain: EGF homology <EGF7>
 F:682-713/Domain: EGF homology <EGF8>
 F:757-788/Domain: EGF homology <EGF9>
 F:795-826/Domain: EGF homology <EGF10>
 F:873-904/Domain: EGF homology <EGF11>
 F:911-942/Domain: EGF homology <EGF12>
 F:949-980/Domain: EGF homology <EGF13>
 F:987-1018/Domain: EGF homology <EGF14>
 F:1025-1056/Domain: EGF homology <EGF15>
 F:1063-1094/Domain: EGF homology <EGF16>
 F:1149-1180/Domain: EGF homology <EGF17>
 F:1187-1218/Domain: EGF homology <EGF18>
 F:1233-1264/Domain: EGF homology <EGF19>
 F:1352-1383/Domain: EGF homology <EGF20>
 F:1391-1425/Domain: EGF homology <EGF>
 F:1491-1498/Domain: ankyrin repeat homology <AN1>
 F:1549-1581/Domain: ankyrin repeat homology <AN2>
 F:1583-2015/Domain: ankyrin repeat homology <AN3>
 F:2016-2048/Domain: ankyrin repeat homology <AN4>
 F:2049-2081/Domain: ankyrin repeat homology <AN5>

Query Match 69.4%; Score 34; DB 2; Length 2531;
 Best Local Similarity 57.1%; Pred. No. 3.2e+02;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNRECT 9
 :|||:|:
 Db 1670 IDNRQCV 1676

RESULT 20
 A40043
 notch protein homolog TAN-1 precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 02-Aug-2002
 C:Accession: A40043
 R:Ellisen, L.W.; Bird, J.; West, D.C.; Soreng, A.L.; Reynolds, T.C.; Smith, S.D.; Sklar, Cell 66, 649-661, 1991
 A:Title: TAN-1, the human homolog of the Drosophila Notch gene, is broken by chromosomal
 A:Reference number: A40043; MID:91347367; PMID:1831692
 A:Accession: A40043
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: mRNA
 A:Residues: 1-2555 <ELL>
 A:Cross-references: GB:M73980
 C:Superfamily: notch protein; ankyrin repeat homology; EGF homology
 F:261-292/Domain: EGF homology <EGF1>
 F:494-525/Domain: EGF homology <EGF1>
 F:987-1018/Domain: EGF homology <EGF2>
 F:1149-1180/Domain: EGF homology <EGF2>
 F:1187-1218/Domain: EGF homology <EGF3>
 F:1233-1264/Domain: EGF homology <EGF3>
 F:1391-1398/Domain: ankyrin repeat homology <AN1>
 F:1491-1498/Domain: ankyrin repeat homology <AN2>
 F:1549-1581/Domain: ankyrin repeat homology <AN3>
 F:1583-2015/Domain: ankyrin repeat homology <AN4>
 F:2016-2048/Domain: ankyrin repeat homology <AN5>

Query Match 69.4%; Score 34; DB 2; Length 2555;
 Best Local Similarity 57.1%; Pred. No. 3.2e+02;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNRECT 9
 :|||:|:
 Db 1680 IDNRQCV 1686

RESULT 21
 KIMST
 thymidine kinase (EC 2.7.1.21), cytosolic - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 04-Dec-1986 #sequence_revision 30-Sep-1993 #text_change 19-Jan-2001
 C:Accession: J01252; A00607; S22599; I58005
 R:Gudas, J.M.; Fridovich-Kell, J.L.; Datta, M.W.; Bryan, J.; Pardee, A.B. Gene 118, 205-216, 1992
 A:Title: Characterization of the murine thymidine kinase-encoding gene and analysis of t
 A:Reference number: J01252; PMID:92380505; PMID:1511894
 A:Accession: J01252
 A:Molecule type: DNA
 A:Residues: 1-233 <GUD>
 A:Cross-references: GB:M68489; MID:9202078; PIDN:AAA0454.1; PID:9202079
 R:Lin, P.F.; Lieberman, H.B.; Yeh, D.B.; Xu, T.; Zhao, S.Y.; Ruddle, F.H. Mol. Cell. Biol. 5, 3149-3156, 1985
 A:Title: Molecular cloning and structural analysis of murine thymidine kinase genomic an
 A:Reference number: A00607; PMID:86310784; PMID:3018504
 A:Accession: A00607
 A:Molecule type: mRNA
 A:Residues: 1-87, 'L', '89-233 <LIN>
 A:Cross-references: GB:M11945; MID:9202072; PIDN:AAA0451.1; PID:9202073
 R:Rotheneder, H.; Grabner, M.; Wintereberger, E. Nucleic Acids Res. 19, 6805-6809, 1991
 A:Title: Presence of regulatory sequences within intron 2 of the mouse thymidine kinase
 A:Reference number: S22599; PMID:92107666; PMID:1762910
 A:Accession: S22599
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-32 <ROT>
 A:Cross-references: EMBL:X60980; MID:G54809; PIDN:CAA43296.1; PID:G54810
 R:Diechty, M.C.; Rauchfuss, H.S.; Iugo, M.H.; Hoxier, J.C. Mutat. Res. 286, 299-307, 1993
 A:Title: Sequence analysis of tka(-)-1 and tkb(+)-1 alleles in L5178Y tk+/- mouse-lympho
 A:Reference number: I58005; MID:93211435; PMID:7681542
 A:Accession: I58005
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-87, 'S', '89-233 <RES>
 A:Cross-references: GB:S57244; MID:9298945; PIDN:AB26001.1; PID:9298946
 C:Comment: Two forms of thymidine kinase, which catalyzes the phosphorylation of thymidi
 tosol. Activity of this cytosolic enzyme is high in proliferating cells and it peaks dur
 C:Genetics: TK1
 A:Introns: 22/3; 33/2; 70/2; 101/3; 131/3; 171/3
 C:Superfamily: thymidine kinase
 C:Keywords: ATP, DNA biosynthesis; nucleic acid binding; P-loop; phosphotransferase
 F:26-33/Region: nucleotide-binding motif A (P-loop)
 F:93-98/Region: nucleotide-binding motif B

Query Match 67.3%; Score 33; DB 1; Length 233;
 Best Local Similarity 55.6%; Pred. No. 51;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGVDNRECT 9
 :|||:|:
 Db 198 AGSDNKNCL 206

RESULT 22
 T27332
 hypothetical protein Y6E2A.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T27332
R/Matthews, L.
submitted to the EMBL Data Library, January 1998
A/Reference number: 220347
A/Accession: T27332
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-307 <WII>
A/Cross-references: EMBL:AL021175; PIDN:CAI5965.1; GSPDB:GN00023; CESP:Y6E2A.1
C/Genetics:
A/Experimental source: clone Y6E2A
A/Map position: 5

Query Match 67.3%; Score 33; DB 2; Length 307;
Best Local Similarity 85.7%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGVDNRE 7
||:||||
Db 195 AGLDNRE 201

RESULT 23
T27333
hypothetical protein Y6E2A.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T27333
R/Matthews, L.
submitted to the EMBL Data Library, January 1998
A/Reference number: 220347
A/Accession: T27333
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-322 <WII>
A/Cross-references: EMBL:AL021175; PIDN:CAI5966.1; GSPDB:GN00023; CESP:Y6E2A.2
C/Genetics:
A/Experimental source: clone Y6E2A
A/Map position: 5
A/Introns: 82/2

Query Match 67.3%; Score 33; DB 2; Length 322;
Best Local Similarity 85.7%; Pred. No. 70;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGVDNRE 7
||:||||
Db 239 AGLDNRE 245

RESULT 24
F83904
hypothetical protein BH2038 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: F83904
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: F83904
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-575 <STO>
A/Cross-references: GB:AP001514; GB:BA000004; NID:g10174613; PIDN:BA05757.1; GSPDB:GN00
C/Genetics:
A/Experimental source: strain C-125
A/Map position: 5

Query Match 67.3%; Score 33; DB 2; Length 575;

Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AGVDNREC 8
||:||||
Db 90 AMLDNREC 97

RESULT 25
T33708
hypothetical protein F58E2.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 08-Sep-2000
C/Accession: T33708
R/Goela, D.; Delenauty, A.
submitted to the EMBL Data Library, October 1998
A/Description: The sequence of C. elegans cosmid F58E2.
A/Reference number: 221390
A/Accession: T33708
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-688 <GOE>
A/Cross-references: EMBL:AF100659; PIDN:AAC68967.1; GSPDB:GN00022; CESP:F58E2.4
A/Experimental source: strain Bristol N2; clone F58E2
C/Genetics:
A/Map position: 4
A/Introns: 228/3; 309/3; 344/2; 602/3
C/Superfamily: Caenorhabditis elegans hypothetical protein F58E2.3

Query Match 67.3%; Score 33; DB 2; Length 688;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

OY 2 GVDNR--ECI 9
||:||||
Db 273 GIDNREIECI 282

RESULT 26
G86450
FSD14.31 protein - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear crease)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C/Accession: G86450
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A6141; MUID:21016719; PMID:11130712
A/Accession: G86450
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-870 <STO>
A/Cross-references: GB:AF005172; NID:g8920629; PIDN:AF81351.1; GSPDB:GN00141
C/Genetics:
A/Map position: 1

Query Match 67.3%; Score 33; DB 2; Length 870;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVDNRECI 9
||:||||
Db 363 GVDNRECI 370

A:Accession: F83636
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-91 <STO>
 A:Cross-references: GB:AE004446; GB:AE004091; NID:g9945886; PIDN:AA03458.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA0068

Query Match 65.3%; Score 32; DB 2; Length 91;
 Best Local Similarity 62.5%; Pred. No. 33;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GVDNRECI 9
 Db 32 GVDNRECI 39

RESULT 32

S69360 retinol-binding protein CRBP, cellular - bovine
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 24-Jul-1998
 C:Accession: S69360

R:Maipelli, G.; Stoppini, M.; Zapponi, M.C.; Foll, C.; Berni, R.
 Eur. J. Biochem. 229, 486-493, 1995

A:Title: Interactions with retinol and retinoids of bovine cellular retinol-binding prot
 A:Reference number: S69360; PMID:95262712; PMID:7744071
 A:Accession: S69360

A:Molecule type: protein
 A:Residues: 1-134 <MAL>

A:Experimental source: bovine testis
 C:Superfamily: myelin P2 protein

C:Keywords: lipid binding; retinol binding; transport protein; vitamin A
 F:1-134/Product: retinol-binding protein, cellular #status experimental <MAT>

Query Match 65.3%; Score 32; DB 2; Length 134;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GVDNRECI 9
 Db 76 GVDNRECI 83

RESULT 33

RJHND retinol-binding protein, cellular - human
 N:Alternate names: cellular retinol-binding protein 2; CRBP
 C:Species: Homo sapiens (man)

C>Date: 30-Jun-1988 #sequence_revision 10-May-1996 #text_change 21-Jul-2000
 C:Accession: S00399; S02300; A40902

R:Nilsson, M.H.L.; Spurr, N.K.; Lundvall, J.; Raek, L.; Peterson, P.A.
 Eur. J. Biochem. 173, 35-44, 1988

A:Title: Human cellular retinol-binding protein gene organization and chromosomal locati
 A:Reference number: S00399; PMID:8818526; PMID:3356192
 A:Accession: S00399

A:Molecule type: DNA
 A:Residues: 1-135 <NIL>

A:Cross-references: EMBL:X07437; NID:G30207; PIDN:CAA0318.1; PID:G4378977
 R:Colantuoni, V.; Correse, R.; Nilsson, M.; Lundvall, J.; Baviak, C.O.; Eriksson, U.; Pet
 Biochem. Biophys. Res. Commun. 130, 431-439, 1985

A:Title: Cloning and sequencing of a full length cDNA corresponding to human cellular re
 A:Reference number: S02300; PMID:85279409; PMID:2992469

A:Accession: S02300
 A:Molecule type: mRNA

A:Residues: 1-135 <COL>
 A:Cross-references: GB:M1433; NID:G190947; PIDN:AAA60257.1; PID:G190948

R:Wei, L.N.; Wertz, J.R.; Goodman, D.S.; Nguyen-Huu, M.C.
 Mol. Endocrinol. 1, 526-534, 1987

A:Title: Cellular retinoic acid- and cellular retinol-binding proteins: complementary de
 A:Reference number: A40902; PMID:91042593; PMID:2856408
 A:Accession: A40902

A:Molecule type: mRNA
 A:Residues: 1-121 <WEI>

A:Cross-references: GB:M63809
 C:Genetics:

A:Gene: GDB:RBP2
 A:Cross-references: GDB:119548; OMIM:180280

A:Map position: 3q21-3qter
 A:Insertion: 25/1; 84/3; 118/3

C:Superfamily: myelin P2 protein
 C:Keywords: lipid binding; transport protein; cellular #status predicted <MAT>
 F:2-135/Product: retinol-binding protein, cellular

Query Match 65.3%; Score 32; DB 1; Length 135;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GVDNRECI 9
 Db 77 GVDNRECI 84

RESULT 34

RJHND retinol-binding protein, cellular - rat
 C:Species: Rattus norvegicus (Norway rat)

C>Date: 05-Apr-1983 #sequence_revision 19-Oct-1995 #text_change 22-Jun-1999
 C:Accession: A29570; A28016; A25227; A92455; A03149

R:Levin, M.S.; Li, E.; Ong, D.E.; Gordon, J.L.
 J. Biol. Chem. 262, 7118-7124, 1987

A:Title: Comparison of the tissue-specific expression and developmental regulation of tw
 A:Reference number: A29570; PMID:87222311; PMID:3584109
 A:Accession: A29570

A:Molecule type: mRNA
 A:Residues: 1-135 <LEV>

A:Cross-references: GB:M19257; GB:002731; NID:G203583; PIDN:AAA40962.1; PID:G203584
 R:Sherman, D.R.; Lloyd, R.S.; Chytil, F.

Proc. Natl. Acad. Sci. U.S.A. 84, 3209-3213, 1987
 A:Title: Rat cellular retinol-binding protein: cDNA sequence and rapid retinol-dependent
 A:Reference number: A28016; PMID:87204107; PMID:3472205

A:Accession: A28016
 A:Molecule type: mRNA

A:Residues: 1-135 <SHR>
 A:Cross-references: GB:M16459; NID:G206590; PIDN:AAA42021.1; PID:G206591

R:Sundelin, J.; Anundt, H.; Tregardh, L.; Eriksson, U.; Lind, P.; Romme, H.; Peterson, P
 J. Biol. Chem. 260, 6488-6493, 1985

A:Title: The primary structure of rat liver cellular retinol-binding protein.
 A:Reference number: A92527; PMID:85207645; PMID:4039728

A:Accession: A92527
 A:Molecule type: protein

A:Residues: 2-135 <SUN>
 A:Experimental source: liver

R:Eriksson, U.; Das, K.; Busch, C.; Nordlinder, H.; Raek, L.; Sundelin, J.; Sallstrom, J
 J. Biol. Chem. 259, 13464-13470, 1984

A:Title: Cellular retinol-binding protein. Quantitation and distribution.
 A:Reference number: A92455; PMID:85030476; PMID:6541654

A:Accession: A92455
 A:Molecule type: protein

A:Residues: 2-51 <ERI>
 A:Experimental source: testis

C:Superfamily: myelin P2 protein
 C:Keywords: lipid binding; transport protein; cellular #status experimental <MAT>
 F:2-135/Product: retinol-binding protein, cellular

Query Match 65.3%; Score 32; DB 1; Length 135;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GVDNRECI 9
 Db 77 GVDNRECI 84

RESULT 35

S16355
 retinol-binding protein I - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 13-Aug-1999
 C:Accession: S16355
 R:Smith, W.C.; Nakshatri, H.; Leroy, P.; Rees, J.; Chambon, P.
 EMBL J. 10, 2223-2230, 1991
 A>Title: A retinoic acid response element is present in the mouse cellular retinol bindi
 A:Reference number: S16355; MUID:91293098; PMID:1648481
 A:Accession: S16355
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-135 <SMI>
 A:Cross-references: GB:X60367; NID:950547; PIDN:CAA42919.1; PID:950548
 C:Superfamily: myelin P2 protein

Query Match 65.3%; Score 32; DB 2; Length 135;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVDNRECT 9
 |||
 Db 77 GIDDRKCM 84

RESULT 36
 E96001
 probable DNA-directed DNA polymerase (EC 2.7.7.7) [imported] - Simorhizobium meliloti (E
 C:Species: Simorhizobium meliloti
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
 C:Accession: E96001
 R:Finan, T.M.; Weidner, S.; Wong, K.; Bhurmesler, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A>Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: E96001
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-213 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC49677.1; PID:G15141164; GSPDB:GNO0167
 A:Experimental source: strain 1021, megaplasmid pSymB
 R:Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 beta, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 283, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kise, E.; Komp, C.; Lelaure,
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A>Title: The composite genome of the legume symbiont Simorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMD21448
 A:Genome: plasmid
 C:Superfamily: Archaeoglobus probable DNA-polymerase
 C:Keywords: nucleotidyltransferase

Query Match 65.3%; Score 32; DB 2; Length 213;
 Best Local Similarity 62.5%; Pred. No. 75;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGVDNREC 8
 |||
 Db 88 AGIDRSEC 95

RESULT 37
 AG3251
 conserved hypothetical protein Atu6193 [imported] - Agrobacterium tumefaciens (strain C5
 C:Species: Agrobacterium tumefaciens
 C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C:Accession: AG3251
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
 erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kutyavah, T.; Levy, R.; Li, M.; McClell

Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Ioo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AG3251
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-243 <KUR>
 A:Cross-references: GB:AB008690; PIDN:AAL6429.1; PID:G17744226; GSPDB:GNO0189
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu6193
 A:Genome: plasmid

Query Match 65.3%; Score 32; DB 2; Length 243;
 Best Local Similarity 62.5%; Pred. No. 85;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGVDNREC 8
 |||
 Db 124 AGADERQC 131

RESULT 38
 A40019
 osteopontin precursor - chicken
 N:Alternate names: bone sialoprotein I, phosphoprotein I, secreted
 C:Species: Gallus gallus (chicken)
 C>Date: 20-Mar-1992 #sequence_revision 15-Aug-1997 #text_change 21-Jul-2000
 C:Accession: I51384; A40019; A37962; S16462
 R:Radfi, K.; Similina, I.; Johnson, E.; Moore, M.A.; Gerstenfeld, L.C.
 Gene 140, 163-169, 1994
 A>Title: Characterization of the chicken osteopontin-encoding gene.
 A:Reference number: I51384; MUID:94192994; PMID:8144023
 A:Accession: I51384
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-264 <RAF>
 A:Cross-references: EMBL:U01844; NID:9404635; PIDN:AAA18584.1; PID:9404636
 R:Castagnola, P.; Bet, P.; Quarto, R.; Gennari, M.; Cancedda, R.
 J. Biol. Chem. 266, 9944-9949, 1991
 A>Title: cDNA cloning and gene expression of chicken osteopontin. Expression of osteopor
 A:Reference number: A40019; MUID:91236779; PMID:2033080
 A:Accession: A40019
 A:Molecule type: mRNA
 A:Residues: 1-215, 'D', 217-234, 'A', 236-239, 'A', 241-264 <CAS>
 A:Cross-references: GB:X56772; NID:958344; PIDN:CAA40091.1; PID:G4493384
 R:Moore, M.A.; Gotch, Y.; Radfi, K.; Gerstenfeld, L.C.
 Biochemistry 30, 2501-2508, 1991
 A>Title: Characterization of a cDNA for chicken osteopontin: expression during bone dev
 A:Reference number: A37962; MUID:91159433; PMID:2001376
 A:Accession: A37962
 A:Molecule type: mRNA
 A:Residues: 1-103, 'G', 105-264 <MOO>
 A:Cross-references: GB:M59182; GB:J05323; NID:9211276; PIDN:AAA62729.1; PID:9211277
 C:Comment: This protein is an acidic glycoprotein rich in aspartic acid, glutamic acid,
 C:Genetics:
 A:Introns: 18/3; 35/3; 70/3; 85/3; 164/3
 C:Superfamily: osteopontin
 C:Keywords: bone; cell binding; extracellular matrix; phosphoprotein; sialoglycoprotein
 F:1-16/Domain: signal sequence; status predicted <SIG>
 F:17-264/Product: osteopontin; status predicted <MAT>
 F:100-106/Region: aspartic acid-rich
 F:138-140/Region: cell attachment (R-G-D) motif
 F:106,109,204,242/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 65.3%; Score 32; DB 1; Length 264;
 Best Local Similarity 66.7%; Pred. No. 92;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGVDNRECT 9

Db 224 AGVDSRESL 232

RESULT 39

F83351

conserved hypothetical protein PA2349 [imported] - Pseudomonas aeruginosa (strain PA01)

C/Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C/Accession: F83351

R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,

; Lory, S.; Olson, M.V.

Nucleotide 406, 959-964, 2000

A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A/Reference number: AB2950; MUID:20437337; PMID:10984043

A/Accession: F83351

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-264 <STO>

A/Cross-references: GB:AE004661; GB:AE004091; NID:g9948382; PIDN:AA05737.1; GSPDB:GN001

A/Experimental source: strain PA01

C/Genetics:

A/Gene: PA2349

C/Superfamily: lipoprotein-28

Query Match 65.3%; Score 32; DB 2; Length 264;

Best Local Similarity 62.5%; Pred. No. 92;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVDNRECT 9

Db 211 GIDNREYV 218

RESULT 40

lipopolysaccharide core biosynthesis protein Waap PA5009 [imported] - Pseudomonas aerugi

C/Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C/Accession: E83020

R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,

; Lory, S.; Olson, M.V.

Nucleotide 406, 959-964, 2000

A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A/Reference number: AB2950; MUID:20437337; PMID:10984043

A/Accession: E83020

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-268 <STO>

A/Cross-references: GB:AE004913; GB:AE004091; NID:g9951286; PIDN:AA08394.1; GSPDB:GN001

A/Experimental source: strain PA01

C/Genetics:

A/Gene: waap, PA5009

Query Match 65.3%; Score 32; DB 2; Length 268;

Best Local Similarity 62.5%; Pred. No. 93;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNREC 8

Db 157 AGVNRHDC 164

Search completed: March 3, 2004, 10:25:25
 Job time : 14.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 09:47:18 ; Search time 10 Seconds
(without alignments)
46.863 Million cell updates/sec

Title: US-09-980-064-2
Perfect score: 49
Sequence: 1 AGVDNRECT 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	49	100.0	VLI_HPV16	P03101 human papill
2	42	85.7	VLI_HPV35	P27232 human papill
3	42	85.7	VLI_HPV52	O05138 human papill
4	41	83.7	VLI_HPV31	P17388 human papill
5	40	81.6	VLI_HPV1	P22163 rhinovirus pap
6	39	79.6	VLI_HPV33	P06416 human papill
7	37	75.5	DSRA_ARCFU	O59109 archaeoglob
8	37	75.5	VLI_HPV58	P26535 human papill
9	34	69.4	RET7_HUMAN	O96R05 homo sapien
10	34	69.4	RET7_MOUSE	O96P05 mus musculu
11	34	69.4	DSVA_DESYA	P45574 desulfovibr
12	34	69.4	NTC2_MOUSE	O35516 mus musculu
13	34	69.4	NTC2_HUMAN	O04721 homo sapien
14	34	69.4	NTC2_RAT	O96W30 rattus norv
15	34	69.4	NTC1_MOUSE	O01705 mus musculu
16	34	69.4	NTC1_RAT	O07008 rattus norv
17	34	69.4	NTC1_HUMAN	P46531 homo sapien
18	33	67.3	KITH_MOUSE	P04184 mus musculu
19	33	67.3	EPAL_HUMAN	P21709 homo sapien
20	32	65.3	PER_METBA	P00202 methanosarc
21	32	65.3	RET1_BOVIN	P02694 bos taurus
22	32	65.3	RET1_HUMAN	P09455 homo sapien
23	32	65.3	RET1_MOUSE	O00915 mus musculu
24	32	65.3	RET1_RAT	P02696 rattus norv
25	32	65.3	OSTP_CHICK	P23498 gallus galli
26	32	65.3	RPNB_SCHPO	P41878 schizosacch
27	32	65.3	VLI_HPV54	P50816 human papil
28	32	65.3	CHIB_EPHAL	P37846 ephedra alt
29	32	65.3	LONI_SPTOL	O04979 spinacia ol
30	32	65.3	GPCR_LYMT	P46023 lymanaea ete
31	32	65.3	MYOG_MOUSE	O64331 mus musculu
32	32	65.3	MYOG_HUMAN	O9um54 homo sapien
33	32	65.3	DYHC_EMENT	P45444 emeticella

34	31	63.3	160	1	YSH2 CAEEL	O09944 caenorhabdi
35	31	63.3	194	1	PYRE CAUCR	O9a810 caulibacter
36	31	63.3	204	1	HISI_STRAM	O99q42 staphylococ
37	31	63.3	265	1	RFAP_ECOLI	P25741 escherichia
38	31	63.3	521	1	CMCH_STRCL	O85728 streptomyce
39	31	63.3	522	1	CP11_ONCMY	O92110 oncorhynch
40	31	63.3	569	1	URE2_HELPFE	O08716 helicobacte
41	31	63.3	590	1	VG29_BPT4	P13337 bacterioph
42	31	63.3	787	1	BPBA_RICPR	O92ce9 rickettsia
43	31	63.3	1175	1	VF19_METVA	O58914 methanococ
44	31	63.3	2164	1	HGAI_PORGI	P59915 porphyromon
45	31	63.3	2437	1	NTC1_BRARE	P46530 brachydanio

ALIGNMENTS

RESULT 1	ID	Query Match	Score	DB 1	Length	511
VLI_HPV16	STANDARD	PRT	531 AA			
AC	P03101					
DT	21-JUL-1986 (Rel. 01, Created)					
DT	21-JUL-1986 (Rel. 01, Last sequence update)					
DT	01-OCT-1996 (Rel. 34, Last annotation update)					
DE	Major capsid protein L1.					
GN	L1.					
OS	Human papillomavirus type 16.					
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;					
OC	Papillomavirus.					
OX	NCBI_TaxID=10581;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=85246220; PubMed=2990099;					
RA	Seedorf K., Krammer G., Durst M., Suhai S., Roweckamp W.G.;					
RT	"Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and variants: a showcase for the molecular evolution of DNA viruses.";					
RL	J. Virol. 66:5714-5725(1992).					
CC	- I- MISCELLANEOUS: HPV16, IN COMPARISON TO HPV TYPES 6 AND 11, IS MORE OFTEN ASSOCIATED WITH MALIGNANT GENITAL CANCERS IN HUMANS.					
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).					
CC	-----					
DR	EMBL; K02718; AAA46943.1; -					
DR	EMBL; M96285; AAA47024.1; -					
DR	EMBL; A06331; CAA00546.1; -					
DR	PIR; A03640; P1WJHS.					
DR	InterPro; IPR002210; PV_capsid_L1.					
DR	InterPro; IPR008975; Viral_cap_coat.					
DR	Pfam; PF00500; late_protein_L1.					
DR	PRINTS; PR00865; HPVCAPSID1.					
DR	ProDom; PD000544; PV_capsid_L1; 1.					
KW	Coat protein; Late protein.					
SQ	SEQUENCE 531 AA; 59554 MW; 583402587093B380 CRC64;					
QY	1 AGVDNRECT 9					
DB	165 AGVDNRECT 173					

```

RESULT 2
V1 HPV35 STANDARD; PRT; 502 AA.
ID V1 HPV35
AC P27232;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 35.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OC NCB1_Taxid=10587;
RN SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
RN SEQUENCE FROM N.A.
RX MEDLINE=92124753; PubMed=1310198;
RA March J.E., Pontsler A.V., Rice S.M., McGraw K.A., Dubenaky T.W.;
RT "The phylogenetic relationship and complete nucleotide sequence of
human papillomavirus type 35.";
RL Virology 186:770-776(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74477; CAA52566.1; -.
DR EMBL; M74117; AAA46972.1; -.
DR PIR; G40824; P1WL35.
DR PIR; S36526; S36526.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; Late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coat protein; Late protein.
FT CONFLICT 94 95 AS -> CL (IN REF. 2).
FT CONFLICT 132 132 S -> L (IN REF. 2).
FT CONFLICT 140 140 G -> GNSG (IN REF. 2).
FT CONFLICT 482 482 K -> R (IN REF. 2).
FT SEQUENCE 502 AA; 56148 MW; 68f6c43508f92267 CRC64;

Query Match 85.7%; Score 42; DB 1; Length 502;
Best Local Similarity 77.8%; Pred. No. 0.55;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
DB 139 SCTDNRCL 147

RESULT 3
V1 HPV52 STANDARD; PRT; 529 AA.
ID V1 HPV52
AC Q05138;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 52.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

```

```

OC Papillomavirus.
OC NCB1_Taxid=10618;
RN SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
RN SEQUENCE OF 333-376 FROM N.A.
RX MEDLINE=92407963; PubMed=1326639;
RA Chan S.Y., Bernard H.U., Ong C.K., Chan S.P., Birgit H., Delius H.;
RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and
variants: a showcase for the molecular evolution of DNA viruses.";
RL Virology 171:306-311(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74481; CAA52590.1; -.
DR EMBL; M96297; AAA47036.1; -.
DR PIR; S36578; S36578.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; Late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coat protein; Late protein.
FT SEQUENCE 529 AA; 59469 MW; 77f7db5c9a664c53 CRC64;

Query Match 85.7%; Score 42; DB 1; Length 529;
Best Local Similarity 75.0%; Pred. No. 0.58;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVDNRECI 9
DB 169 GIDNRECL 176

RESULT 4
V1 HPV31 STANDARD; PRT; 504 AA.
ID V1 HPV31
AC P17388;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 31.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OC NCB1_Taxid=10585;
RN SEQUENCE FROM N.A.
RX MEDLINE=89299478; PubMed=2545036;
RA Goldsborough M.D., Disilvestre D., Temple G.F., Iorincz A.T.;
RT "Nucleotide sequence of human papillomavirus type 31: a cervical
neoplasia-associated virus.";
RL Virology 171:306-311(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL; J04353; AAA46956.1; -.
 DR PIR; G32444; P1WL31.
 DR InterPro; IPR002210; PV_capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsid.L1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 DR Coat protein; Late protein.
 SQ SEQUENCE 504 AA; 56352 MW; B45A306A6B3AB9D2 CRC64;

Query March 83.7%; Score 41; DB 1; Length 504;
 Best Local Similarity 87.5%; Pred. No. 0.88;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVNDRECT 9
 DB 141 GTDRECT 148

RESULT 5
 VLI_RHPV1 STANDARD; PRT; 501 AA.

AC P2163;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Major capsid protein L1.

GN L1.
 OS Rhesus papillomavirus type 1 (Rhpv 1).
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCB1_TaxID=10570;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91135018; PubMed=1847267;
 RA Ostrrow R.S., Labresh K.V., Faras A.J.;
 RT "Characterization of the complete RHPV 1 genomic sequence and an
 RL integration locus from a metastatic tumor.";
 RL Virology 181:424-429(1991).
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; M60184; AAA79318.1; -.
 DR PIR; H3503; P1WLR1.
 DR InterPro; IPR002210; PV_capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsid.L1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 DR Coat protein; Late protein.
 SQ SEQUENCE 501 AA; 55635 MW; 6E808D8F130E440A CRC64;

Query March 81.6%; Score 40; DB 1; Length 501;
 Best Local Similarity 75.0%; Pred. No. 1.4;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVNDRECT 9
 DB 139 GADNRECT 146

RESULT 6
 VLI_HPV33 STANDARD; PRT; 499 AA.
 ID VLI_HPV33
 AC P0616;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Major capsid protein L1.

GN L1.
 OS Human papillomavirus type 33.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCB1_TaxID=10586;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86200464; PubMed=3009902;
 RA Cole S.T., Strecek R.E.;
 RT "Genome organization and nucleotide sequence of human papillomavirus
 RT type 33, which is associated with cervical cancer.";
 RL J. Virol. 58:991-995(1986).
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M12732; AAA46964.1; -.
 DR PIR; A03641; P1WL33.
 DR InterPro; IPR002210; PV_capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsid.L1.
 DR ProDom; PD000544; PV_capsid_L1; 1.
 DR Coat protein; Late protein.
 KM SEQUENCE 499 AA; 55902 MW; 283474DA93DE01D CRC64;

Query March 79.6%; Score 39; DB 1; Length 499;
 Best Local Similarity 75.0%; Pred. No. 2.3;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVNDRECT 9
 DB 141 GADNRECT 148

RESULT 7

ID DSR_ARCFU STANDARD; PRT; 417 AA.
 AC Q59109;

DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Sulfite reductase, dissimilatory-type alpha subunit (EC 1.8.99.3)
 DE (Hydrogensulfite reductase alpha subunit).
 GN DSR_A OR AF0423.

OS Archaeoglobus fulgidus.
 OC Archaea; Buryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCB1_TaxID=2234;

RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-36; 64-70; 81-98; 123-140;
 RP 343-354; 361-367 AND 391-412.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=94014979; PubMed=7691984;
 RA Dahl C., Kredich N.M., Deutzmann R., Truper H.G.;
 RT "Dissimilatory sulfite reductase from Archaeoglobus fulgidus:
 RT physico-chemical properties of the enzyme and cloning, sequencing
 RT and analysis of the reductase genes.";
 RL J. Gen. Microbiol. 139:1817-1828(1993).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kierlavage A.R., Graham D.E., Kyrpides N.C.,

```

RA Pfleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirtane S.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodex A., Zhou L.,
RA Overbeek R., Cocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Moese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RT Nature 390:364-370(1997).
CC -1- FUNCTION: Catalyzes the reduction of sulfite to sulfide. This is
CC the terminal oxidation reaction in sulfate respiration. Extremely
CC thermophilic, no activity towards methylololigen was found below
CC 55 degrees Celsius.
CC -1- CATALYTIC ACTIVITY: (O(3).S.S.SO(3)) (2-) + acceptor + 2 H(2)O +
CC OH(-) = 3 (HSO(3)) (-) + reduced acceptor.
CC -1- COFACTOR: Binds 1 4Fe-4S iron-sulfur cluster and 2 sirohemes per
CC subunit.
CC -1- SUBUNIT: Heterotetramer of two alpha and two beta subunits.
CC -1- SUBCELLULAR LOCATION: Although the protein complex is found in the
CC soluble fraction it may be membrane-associated in vivo.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M95634; AAB17213.1; -
DR EMBL; AE001075; AAB90812.1; -
DR PIR; G69302; G69302.
DR TIGR; AF0423; -
DR InterPro; IPR006067; Nix_Sir_4Fe4S.
DR Pfam; PF01077; Nix_Sir; 1.
DR Oxioreductase; Heme; Iron-sulfur; 4Fe-4S; Complete proteome.
KW INIT_MET
FT METAL 175 175 IRON (HEME AXIAL LIGAND) (POTENTIAL).
FT METAL 181 181 IRON (HEME AXIAL LIGAND) (POTENTIAL).
FT METAL 219 219 IRON (HEME AXIAL LIGAND) (POTENTIAL).
FT METAL 223 223 IRON (HEME AXIAL LIGAND) (POTENTIAL).
FT METAL 226 226 IRON-SULFUR (4Fe-4S) (POTENTIAL).
FT METAL 265 265 IRON-SULFUR (4Fe-4S) (POTENTIAL).
FT METAL 285 285 IRON-SULFUR (4Fe-4S) (POTENTIAL).
FT METAL 288 288 IRON-SULFUR (4Fe-4S) (POTENTIAL).
FT METAL 291 291 IRON-SULFUR (4Fe-4S) (POTENTIAL).
SQ SEQUENCE 417 AA; 47393 MW; 18E70E7B3F767373 CRC64;
Query Match 75.5%; Score 37; DB 1; Length 417;
Best Local Similarity 71.4%; Pred. No. 4.8;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 3 VDRECT 9
DB 280 IDNRECV 286

```

```

RX MEDLINE=92024102; PubMed=1656594;
RA Kiriti Y., Iwamoto S., Matsukura T.;
RT "Human papillomavirus type 58 DNA sequence."
RT Virology 185:424-427(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90400; BAA31851.1; -
DR PIR; G36779; P1M58.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coat protein; Late protein.
SQ SEQUENCE 524 AA; 59038 MW; 5DEC95B2878A26DF CRC64;
Query Match 75.5%; Score 37; DB 1; Length 524;
Best Local Similarity 75.0%; Pred. No. 6.1;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 GVDRECT 9
DB 167 GSDNRECL 174

```

```

RESULT 8
ID VLI HPV58 STANDARD; PRT; 524 AA.
AC P26535;
DT 01-AUG-1992 (rel. 23, Created)
DT 01-AUG-1992 (rel. 23, Last sequence update)
DT 01-OCT-1996 (rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 58.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxId=10598;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RESULT 9
ID RET7 HUMAN STANDARD; PRT; 134 AA.
AC Q96R05;
DT 28-FEB-2003 (rel. 41, Created)
DT 28-FEB-2003 (rel. 41, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE Retinol-binding protein IV, cellular (CRBP-IV) (Retinoid binding
DE protein 7).
GN RBP7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND X-RAY CRYSTALLOGRAPHY (2.0
RP ANGSTROMS).
RX MEDLINE=22289282; PubMed=12177003;
RA Follis C., Calderone V., Ramazzina I., Zanotti G., Berni R.;
RT "Ligand binding and structural analysis of a human putative cellular
RT retinol-binding protein."
RL J. Biol. Chem. 277:41970-41977(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Guo J.H., She X.Y., Dai F.Y., Yu L.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Intracellular transport of retinol.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed primarily in kidney, heart and
CC transverse colon.
CC -1- SIMILARITY: Belongs to the fatty-acid binding protein (FABP)
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; AY145438; AAN61071.1; -.
DR EMBL; AF399927; AAK85409.1; -.
DR PDB; 1LPJ; 14-JAN-03.
DR InterPro; IPR000463; Fatty acid BP.
DR InterPro; IPR000566; Lipocin_cytfabp.
DR Pfam; PF00061; lipocalin; 1.
DR PRINTS; PR00178; FATTYACIDBP.
DR PROSITE; PS00214; FAPB; 1.
DR Vitamin A; Retinol-binding; Transport; 3D-structure.
SQ SEQUENCE 134 AA; 15536 MW; B152C8B3D95B20BD CRC64;

Query Match
Best Local Similarity 69.4%; Score 34; DB 1; Length 134;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVDNREC 8
Db 77 GLDNKRC 83

RESULT 10
RETI MOUSE STANDARD; PRT; 134 AA.
ID RETI_MOUSE
AC O9EPK5; O9CTKO;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Retinol-binding protein IV, cellular (CRBP-IV) (Retinoid binding
DE protein 7).
GN RBP7.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCB1_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=129, C57BL/6J, and C57BL/6J;
RC MEDLINE=20481916; PubMed=11027338;
RA Conforti L., Tarillon A., Mack T.G.A., Mi W., Buckmaster E.A.,
RA Wagner D., Perry V.H., Coleman M.P.;
RA "A Ufd3/D4Colele chimeric protein and overexpression of Rbp7 in the
RT slow Wallerian degeneration (Wlds) mouse."
RL Proc. Natl. Acad. Sci. U.S.A. 97:11377-11382 (2000).
[2]
RN SEQUENCE OF 59-134 FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Embryo;
RC MEDLINE=21085660; PubMed=11217851;
RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aikawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kaaukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Koehliwa H.,
RA Kuehl P., Lewis S., Matsumoto Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baladevall R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gutierrez S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzaroli J., Mombere P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlschki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690 (2001).
CC -1- FUNCTION: Intracellular transport of retinol (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- TISSUE SPECIFICITY: Highly expressed in white adipose tissue and
CC mammary gland.
CC -1- SIMILARITY: Belongs to the fatty-acid binding protein (FABP)
CC family.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announce/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; AF260927; AAG38491.1; -.
DR EMBL; AF260923; AAG17284.1; -.
DR EMBL; AK003307; BAB22705.1; -.
DR HSSP; P02696; ICRB.
DR MED; MGI:1890409; Rbp7.
DR GO; GO:0005501; F:retinoid binding; IDA.
DR InterPro; IPR000463; Fatty acid BP.
DR InterPro; IPR000566; Lipocin_cytfabp.
DR Pfam; PF00061; lipocalin; 1.
DR PRINTS; PR00178; FATTYACIDBP.
DR PROSITE; PS00214; FAPB; 1.
DR Vitamin A; Retinol-binding; Transport.
SQ SEQUENCE 134 AA; 15427 MW; 04D1A591AD7263B6 CRC64;

Query Match
Best Local Similarity 71.4%; Score 34; DB 1; Length 134;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVDNREC 8
Db 77 GLDNKRC 83

RESULT 11
IDV DESVH STANDARD; PRT; 436 AA.
IDV DESVH
AC P45574; Q46581;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sulfite reductase, dissimilatory-type alpha subunit (EC 1.8.99.3)
DE (Desulfovibrio alpha subunit) (Hydrogensulfite reductase alpha
DE subunit).
GN DESVH.
OS Desulfovibrio vulgaris (strain Hildenborough).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCB1_TaxID=882;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95193998; PubMed=7887608;
RA Karthoff-Schweizer R.R., Huber D.P.W., Voordouw G.;
RT "Conservation of the genes for dissimilatory sulfite reductase from
RT Desulfovibrio vulgaris and Archaeoglobus fulgidus allows their
RT detection by PCR."
RL Appl. Environ. Microbiol. 61:290-296 (1995).
[2]
RN SEQUENCE OF 1-25.
RP MEDLINE=92209491; PubMed=1555572;
RX Pierik A.J., Duyvis M.G., van Helvoort J.M.L.M., Wolbert R.B.G.,
RA Hagen W.R.;
RT "The third subunit of desulfovibrio-type dissimilatory sulfite
RT reductases."
RL Eur. J. Biochem. 205:111-115 (1992).
CC -1- FUNCTION: Catalyzes the reduction of sulfite to sulfide. This is
CC the terminal oxidation reaction in sulfate respiration, a process
CC catalyzed by the sulfate-reducing bacteria.
CC -1- CATALYTIC ACTIVITY: (O(3)S.S.SO(3))^(2-) + acceptor + 2 H(2)O +
CC OH(-) = 3 (HSO(3))^(1-) + reduced acceptor.
CC -1- COFACTOR: Binds 1 4Fe-4S iron-sulfur cluster and 2 sirohemes per
CC subunit.
CC -1- SUBUNIT: Heterohexamer of two alpha, two beta and two gamma
CC subunits.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: U06723; AAA70107.1; -.
 CC PIR: S21197; S21197.
 CC InterPro: IPR006067; Nlr_Slr_4Fe4S.
 CC Pfam: PF01077; Nlr_Slr; 1.
 CC DR Oxidoreductase; Heme; Iron-sulfur; 4Fe-4S.
 CC
 CC INIT MET 0
 CC METAL 176 176 IRON (HEME AXIAL LIGAND) (POTENTIAL).
 CC METAL 182 182 IRON (HEME AXIAL LIGAND) (POTENTIAL).
 CC METAL 220 220 IRON (HEME AXIAL LIGAND) (POTENTIAL).
 CC METAL 224 224 IRON (HEME AXIAL LIGAND) (POTENTIAL).
 CC METAL 283 283 IRON-SULFUR (4FE-4S) (POTENTIAL).
 CC METAL 302 302 IRON-SULFUR (4FE-4S) (POTENTIAL).
 CC METAL 305 305 IRON-SULFUR (4FE-4S) (POTENTIAL).
 CC METAL 308 308 IRON-SULFUR (4FE-4S) (POTENTIAL).
 CC METAL 21 21 S -> V (IN REF. 2).
 CC CONFLICT 25 25 S -> K (IN REF. 2).
 CC FT CONFLICT 25 25 Q -> K (IN REF. 2).
 CC SQ SEQUENCE 436 AA; 48959 MM; 7AA5895FA4F31910 CRC64;

Query Match 69.4%; Score 34; DB 1; Length 436;
 Best Local Similarity: 57.1%; Pred. No. 21;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNRECT 9
 Db 297 IDMKCV 303

RESULT 12
 NTG2_MOUSE STANDARD; PRT; 2470 AA.
 ID NTG2_MOUSE STANDARD; PRT; 2470 AA.
 AC 035516; 006008; 060941;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neutrogenic locus notch homolog protein 2 precursor (Notch 2) (Notch
 B).
 GN NOTCH2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Thymus;
 RA Hamada Y., Higuchi M., Teujimoto Y.;
 RT "Complete amino acid sequence and multiform transcripts encoded by a
 RT single copy of mouse Notch2 gene."
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 316-1518 FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Embryo;
 RX MEDLINE=93178563; PubMed=8440332;
 RA Lardelli M., Lendahl U.;
 RT "Notch A and Notch B-two mouse Notch homologues coexpressed in a
 RT wide variety of tissues."
 RL Exp. Cell Res. 204:364-372(1993).
 RN [3]
 RP SEQUENCE OF 1765-2153 FROM N.A.
 RX MEDLINE=97075110; PubMed=8917536;
 RA Milner L.A., Bigas A., Kopan R., Brashem-Stein C., Bernstein I.D.,
 RT Martin D.I.;
 RT "Inhibition of granulocytic differentiation by mNotch1."
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13014-13019(1996).
 RN [4]
 RP FUNCTION.

RX MEDLINE=99396706; PubMed=10393120;
 RA Hamada Y., Kadokawa Y., Okabe M., Ikawa M., Coleman J.R.,
 RA Teujimoto Y.;
 RT "Mutation in ankyrin repeats of the mouse Notch2 gene induces early
 RT embryonic lethality."
 RL Development 126:3415-3424(1999).
 RN [5]
 RP DEVELOPMENTAL STAGE, AND ALTERNATIVE SPLICING.
 RX MEDLINE=9533893; PubMed=7609614;
 RA Higuchi M., Kiyama H., Hayakawa T., Hamada Y., Teujimoto Y.;
 RT "Differential expression of Notch1 and Notch2 in developing and adult
 RT mouse brain."
 RL Brain Res. Mol. Brain Res. 29:263-272(1995).
 RN [6]
 RP POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS OF MET-1699.
 RX MEDLINE=21374376; PubMed=11518718;
 RA Saxena M.T., Schroeter E.H., Mumm J.S., Kopan R.;
 RT "Murine notch homologs (NL-4) undergo presenilin-dependent
 RT proteolysis."
 RL J. Biol. Chem. 276:40268-40273(2001).
 RN [7]
 RP POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS OF MET-1699.
 RX MEDLINE=21374376; PubMed=11459941;
 RA Mizutani T., Taniguchi Y., Aoki T., Hashimoto N., Honjo T.;
 RT "Conservation of the biochemical mechanisms of signal transduction
 RT among mammalian Notch family members."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031(2001).
 CC -1- FUNCTION: Functions as a receptor for membrane-bound ligands
 CC Jagged1, Jagged2 and Delta to regulate cell-fate determination.
 CC Upon ligand activation through the released cell-intracellular
 CC domain (NICD) it forms a transcriptional activator complex with
 CC RBP-J kappa and activates genes of the enhancer of split locus.
 CC Affects the implementation of differentiation, proliferation and
 CC apoptotic programs (By similarity). May play an essential role in
 CC postimplantation development, probably in some aspect of cell
 CC specification and/or differentiation.
 CC -1- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
 CC terminal fragment N(IEC) which are probably linked by disulfide
 CC bonds.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Following
 CC proteolytic processing NICD is translocated to the nucleus.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=035516-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=035516-2; Sequence=VSP 001405;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Expressed in the brain, liver, kidney,
 CC neuroepithelia, somites, optic vesicles and branchial arches, but
 CC not heart.
 CC -1- DEVELOPMENTAL STAGE: Expressed in the embryonic ventricular zone,
 CC the postnatal ependymal cells, and the choroid plexus throughout
 CC embryonic and postnatal development.
 CC -1- PTM: Synthesized in the endoplasmic reticulum as an inactive form
 CC which is proteolytically cleaved by a furin-like convertase to yield
 CC trans-Golgi network before it reaches the plasma membrane to yield
 CC an active, ligand-accessible form. Cleavage results in a C-
 CC terminal fragment N(TM) and a N-terminal fragment N(IEC). Following
 CC ligand binding, it is cleaved by TNF-alpha converting enzyme
 CC (TACE) to yield a membrane-associated intermediate fragment called
 CC notch extracellular truncation (NEXTr). This fragment is then
 CC cleaved by presenilin dependent gamma-secretase to release a
 CC notch-derived peptide containing the intracellular domain (NICD)
 CC from the membrane.
 CC -1- PTM: Phosphorylated.
 CC -1- SIMILARITY: Belongs to the NOTCH family.
 CC -1- SIMILARITY: Contains 35 EGF-like domains.
 CC -1- SIMILARITY: Contains 2 Lin/Notch repeats.
 CC -1- SIMILARITY: Contains 6 ANK repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

FT	DOMAIN	793	829		EGF-LIKE 21,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	831	869		EGF-LIKE 22,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	871	907		EGF-LIKE 23,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	909	945		EGF-LIKE 24,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	947	983		EGF-LIKE 25,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	985	1021		EGF-LIKE 26,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	1023	1059		EGF-LIKE 27,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	1061	1097		EGF-LIKE 28,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	1099	1145		EGF-LIKE 29,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	1147	1183		EGF-LIKE 30,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	1185	1221		EGF-LIKE 31,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	1223	1260		EGF-LIKE 32,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	1262	1300		EGF-LIKE 33,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	1302	1345		EGF-LIKE 34,	CALCIUM-BINDING (POTENTIAL)	.
FT	DOMAIN	1372	1410		EGF-LIKE 35,	CALCIUM-BINDING (POTENTIAL)	.
FT	REPEAT	1418	1454		LIN/NOTCH 1.	LIN/NOTCH 2.	.
FT	REPEAT	1501	1533		LIN/NOTCH 2.	LIN/NOTCH 2.	.
FT	REPEAT	1825	1869		ANK 1.	ANK 1.	.

Query Match Best Local Similarity 69.4%; Score 34; DB 1; Length 2470;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 VDRNCCI 9
Db 1627 IDNRQCV 1633

```

RESULT 13
NTC2_HUMAN          STANDARD;      PRT:   2471 AA.
AC    Q04721; Q99734; Q9H240;
AD    28-FEB-2003 (Rel. 41, Created)
DT    28-FEB-2003 (Rel. 41, Last sequence update)
DT    10-OCT-2003 (Rel. 42, Last annotation update)
DE    Neurogenic locus notch homolog protein 2 precursor (Notch 2) (hn2).
GN    NOTCH2.
OS    Homo sapiens (Human).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX    NCBI_TaxID=9606;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Brain;
RC    Blaumüller C.M., Mann R.S.;
RA    "Complete human notch 2 (hn2) cDNA sequence.";
RT    Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RL    [2]
RN    SEQUENCE FROM N.A.
RC    TISSUE=Breast tumor;
RC    Correa R.G., Camargo A.A., Moreira F.S., Simpson A.J.G.;
RA    "Human Notch2, a novel member of cell-fate determining NOTCH
RT    family.";
RL    Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN    [3]
RP    SEQUENCE OF 967-1229 FROM N.A.
RC    TISSUE=T-cell;
RA    Lemasson I., Devaux C., Meunard J.M.;
RT    "Partial sequence of EGF-like repeat domain of human Notch2 mRNA.";
RL    Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN    [4]
RP    SEQUENCE OF 1810-2447 FROM N.A.
RC    TISSUE=Brain;
RC    MEDLINE=93265135; PubMed=1303360;
RA    Stefani S., Blaumüller C.M., Redhead N.J., Hill R.E.,
RT    Artavanis-Tsakonas S.;
RT    "Human homologs of a Drosophila enhancer of split gene product define
RL    a novel family of nuclear proteins.";
RN    Nat. Genet. 2:119-127(1992).
RN    [5]
RP    POST-TRANSLATIONAL PROCESSING.
RC    MEDLINE=97386453; PubMed=9244302;
RA    Blaumüller C.M., Qi H., Zagouras P., Artavanis-Tsakonas S.;

```

```

RT "intracellular cleavage of Notch leads to a heterodimeric receptor on
RT the plasma membrane."
RT Cell 90:281-291(1997).
RN [6]
RP IDENTIFICATION OF LIGANDS.
RX MEDLINE=99180765; PubMed=10079256;
RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.L.,
RT Human ligands of the Notch receptor."
RT Am. J. Pathol. 154:785-794(1999).
CC -1- FUNCTION: Functions as a receptor for membrane-bound ligands
CC Jagged1, Jagged2 and Delta1 to regulate cell-fate determination.
CC Upon ligand activation through the released cell-surface Notch
CC domain (NICD) it forms a transcriptional activator complex with
CC RBP-J kappa and activates genes of the enhancer of split locus.
CC Affects the implementation of differentiation, proliferation and
CC apoptotic programs (By similarity).
CC SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
CC terminal fragment N(EG) which are probably linked by disulfide
CC bonds (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Following
CC proteolytic processing NICD is translocated to the nucleus.
CC -1- TISSUE SPECIFICITY: Expressed in the brain, heart, kidney, lung,
CC skeletal muscle and liver.
CC -1- PTM: Synthesized in the endoplasmic reticulum as an inactive form
CC which is proteolytically cleaved by a furin-like convertase in the
CC trans-Golgi network before it reaches the plasma membrane to yield
CC an active, ligand-accessible form. Cleavage results in a C-
CC terminal fragment N(TM) and a N-terminal fragment N(EG). Following
CC ligand binding, it is cleaved by TNF-alpha converting enzyme
CC (TACE) to yield a membrane-associated intermediate fragment called
CC notch extracellular truncation (NEXT). This fragment is then
CC cleaved by presenilin dependent gamma-secretase to release a
CC notch-derived peptide containing the intracellular domain (NICD)
CC from the membrane (By similarity).
CC -1- PTM: Phosphorylated (By similarity).
CC -1- SIMILARITY: Belongs to the NOTCH family.
CC -1- SIMILARITY: Contains 35 EGF-like domains.
CC -1- SIMILARITY: Contains 2 Lin/Notch repeats.
CC -1- SIMILARITY: Contains 6 ANK repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF08601; AAA36377.2; -
DR EMBL: AF315356; MAG37073.1; -
DR EMBL: U77493; AAB19224.1; -
DR HSSP: P00740; 1EDM.
DR Genew; HGNC:7882; NOTCH2.
DR MIM: 600275; -
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000152; Asx_hydroxyl_5.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_II.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR008297; Notch.
DR InterPro: IPR000800; Notch_dom.
DR Pfam: PF00023; ank; 6.
DR Pfam: PF00066; notch; 2.
DR PIRSF: PIRSF002279; Notch; 1.
DR PRINTS: PRO0010; EGFBLCOO.
DR PRINTS: PRO0011; EGFFLAMININ.
DR PRINTS: PRO1452; NOTCH.
DR SMART; SM00248; ANK; 6.
DR SMART; SM00179; EGF_CA; 23.

```

```

DR SMART; SM00004; NL; 2.
DR PROSITE; PSS0287; ANK_REPEAT; 1.
DR PROSITE; PSS0088; ANK_REPEAT; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 22.
DR PROSITE; PS00022; EGF_1; 34.
DR PROSITE; PS01186; EGF_2; 29.
DR PROSITE; PS00026; EGF_3; 35.
DR PROSITE; PS01187; EGF_CA; 22.
KW Receptor; Transcription regulation; Activator; Differentiation;
KW Developmental protein; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Glycoprotein;
KW SIGNAL
FT 1 25
FT CHAIN
FT 1666 2471
FT CHAIN
FT 1697 2471
FT CHAIN
FT 26 1677
FT TRANSMEM
FT 1678 1698
FT 1699 2471
FT 26 63
FT 64 102
FT 105 143
FT 144 180
FT 182 219
FT 221 258
FT 260 296
FT 298 336
FT 338 374
FT 375 413
FT 415 454
FT 456 492
FT 494 530
FT 532 568
FT 570 605
FT 607 643
FT 645 680
FT 682 718
FT 720 755
FT 757 793
FT 795 831
FT 871 909
FT 911 947
FT 949 985
FT 987 1023
FT 1025 1061
FT 1063 1099
FT 1101 1147
FT 1149 1185
FT 1187 1223
FT 1225 1262
FT 1264 1302
FT 1304 1343
FT 1374 1412
FT 1420 1456
FT 1456 1503
FT 1503 1535
FT 1535 1571
FT 1571 1607
FT 1607 1643
FT 1643 1679
FT 1679 1715
FT 1715 1751
FT 1751 1787
FT 1787 1823
FT 1823 1859
FT 1859 1895
FT 1895 1931
FT 1931 1967
FT 1967 2003
FT 2003 2039
FT 2039 2075
FT 2075 2111
FT 2111 2147
FT 2147 2183
FT 2183 2219
FT 2219 2255
FT 2255 2291
FT 2291 2327
FT 2327 2363
FT 2363 2399
FT 2399 2435
FT 2435 2471
FT 2471 2507
FT 2507 2543
FT 2543 2579
FT 2579 2615
FT 2615 2651
FT 2651 2687
FT 2687 2723
FT 2723 2759
FT 2759 2795
FT 2795 2831
FT 2831 2867
FT 2867 2903
FT 2903 2939
FT 2939 2975
FT 2975 3011
FT 3011 3047
FT 3047 3083
FT 3083 3119
FT 3119 3155
FT 3155 3191
FT 3191 3227
FT 3227 3263
FT 3263 3299
FT 3299 3335
FT 3335 3371
FT 3371 3407
FT 3407 3443
FT 3443 3479
FT 3479 3515
FT 3515 3551
FT 3551 3587
FT 3587 3623
FT 3623 3659
FT 3659 3695
FT 3695 3731
FT 3731 3767
FT 3767 3803
FT 3803 3839
FT 3839 3875
FT 3875 3911
FT 3911 3947
FT 3947 3983
FT 3983 4019
FT 4019 4055
FT 4055 4091
FT 4091 4127
FT 4127 4163
FT 4163 4199
FT 4199 4235
FT 4235 4271
FT 4271 4307
FT 4307 4343
FT 4343 4379
FT 4379 4415
FT 4415 4451
FT 4451 4487
FT 4487 4523
FT 4523 4559
FT 4559 4595
FT 4595 4631
FT 4631 4667
FT 4667 4703
FT 4703 4739
FT 4739 4775
FT 4775 4811
FT 4811 4847
FT 4847 4883
FT 4883 4919
FT 4919 4955
FT 4955 4991
FT 4991 5027
FT 5027 5063
FT 5063 5099
FT 5099 5135
FT 5135 5171
FT 5171 5207
FT 5207 5243
FT 5243 5279
FT 5279 5315
FT 5315 5351
FT 5351 5387
FT 5387 5423
FT 5423 5459
FT 5459 5495
FT 5495 5531
FT 5531 5567
FT 5567 5603
FT 5603 5639
FT 5639 5675
FT 5675 5711
FT 5711 5747
FT 5747 5783
FT 5783 5819
FT 5819 5855
FT 5855 5891
FT 5891 5927
FT 5927 5963
FT 5963 5999
FT 5999 6035
FT 6035 6071
FT 6071 6107
FT 6107 6143
FT 6143 6179
FT 6179 6215
FT 6215 6251
FT 6251 6287
FT 6287 6323
FT 6323 6359
FT 6359 6395
FT 6395 6431
FT 6431 6467
FT 6467 6503
FT 6503 6539
FT 6539 6575
FT 6575 6611
FT 6611 6647
FT 6647 6683
FT 6683 6719
FT 6719 6755
FT 6755 6791
FT 6791 6827
FT 6827 6863
FT 6863 6899
FT 6899 6935
FT 6935 6971
FT 6971 7007
FT 7007 7043
FT 7043 7079
FT 7079 7115
FT 7115 7151
FT 7151 7187
FT 7187 7223
FT 7223 7259
FT 7259 7295
FT 7295 7331
FT 7331 7367
FT 7367 7403
FT 7403 7439
FT 7439 7475
FT 7475 7511
FT 7511 7547
FT 7547 7583
FT 7583 7619
FT 7619 7655
FT 7655 7691
FT 7691 7727
FT 7727 7763
FT 7763 7799
FT 7799 7835
FT 7835 7871
FT 7871 7907
FT 7907 7943
FT 7943 7979
FT 7979 8015
FT 8015 8051
FT 8051 8087
FT 8087 8123
FT 8123 8159
FT 8159 8195
FT 8195 8231
FT 8231 8267
FT 8267 8303
FT 8303 8339
FT 8339 8375
FT 8375 8411
FT 8411 8447
FT 8447 8483
FT 8483 8519
FT 8519 8555
FT 8555 8591
FT 8591 8627
FT 8627 8663
FT 8663 8699
FT 8699 8735
FT 8735 8771
FT 8771 8807
FT 8807 8843
FT 8843 8879
FT 8879 8915
FT 8915 8951
FT 8951 8987
FT 8987 9023
FT 9023 9059
FT 9059 9095
FT 9095 9131
FT 9131 9167
FT 9167 9203
FT 9203 9239
FT 9239 9275
FT 9275 9311
FT 9311 9347
FT 9347 9383
FT 9383 9419
FT 9419 9455
FT 9455 9491
FT 9491 9527
FT 9527 9563
FT 9563 9599
FT 9599 9635
FT 9635 9671
FT 9671 9707
FT 9707 9743
FT 9743 9779
FT 9779 9815
FT 9815 9851
FT 9851 9887
FT 9887 9923
FT 9923 9959
FT 9959 9995
FT 10000 10036
FT 10036 10072
FT 10072 10108
FT 10108 10144
FT 10144 10180
FT 10180 10216
FT 10216 10252
FT 10252 10288
FT 10288 10324
FT 10324 10360
FT 10360 10396
FT 10396 10432
FT 10432 10468
FT 10468 10504
FT 10504 10540
FT 10540 10576
FT 10576 10612
FT 10612 10648
FT 10648 10684
FT 10684 10720
FT 10720 10756
FT 10756 10792
FT 10792 10828
FT 10828 10864
FT 10864 10900
FT 10900 10936
FT 10936 10972
FT 10972 11008
FT 11008 11044
FT 11044 11080
FT 11080 11116
FT 11116 11152
FT 11152 11188
FT 11188 11224
FT 11224 11260
FT 11260 11296
FT 11296 11332
FT 11332 11368
FT 11368 11404
FT 11404 11440
FT 11440 11476
FT 11476 11512
FT 11512 11548
FT 11548 11584
FT 11584 11620
FT 11620 11656
FT 11656 11692
FT 11692 11728
FT 11728 11764
FT 11764 11800
FT 11800 11836
FT 11836 11872
FT 11872 11908
FT 11908 11944
FT 11944 11980
FT 11980 12016
FT 12016 12052
FT 12052 12088
FT 12088 12124
FT 12124 12160
FT 12160 12196
FT 12196 12232
FT 12232 12268
FT 12268 12304
FT 12304 12340
FT 12340 12376
FT 12376 12412
FT 12412 12448
FT 12448 12484
FT 12484 12520
FT 12520 12556
FT 12556 12592
FT 12592 12628
FT 12628 12664
FT 12664 12700
FT 12700 12736
FT 12736 12772
FT 12772 12808
FT 12808 12844
FT 12844 12880
FT 12880 12916
FT 12916 12952
FT 12952 12988
FT 12988 13024
FT 13024 13060
FT 13060 13096
FT 13096 13132
FT 13132 13168
FT 13168 13204
FT 13204 13240
FT 13240 13276
FT 13276 13312
FT 13312 13348
FT 13348 13384
FT 13384 13420
FT 13420 13456
FT 13456 13492
FT 13492 13528
FT 13528 13564
FT 13564 13600
FT 13600 13636
FT 13636 13672
FT 13672 13708
FT 13708 13744
FT 13744 13780
FT 13780 13816
FT 13816 13852
FT 13852 13888
FT 13888 13924
FT 13924 13960
FT 13960 14000
FT 14000 14040
FT 14040 14080
FT 14080 14120
FT 14120 14160
FT 14160 14200
FT 14200 14240
FT 14240 14280
FT 14280 14320
FT 14320 14360
FT 14360 14400
FT 14400 14440
FT 14440 14480
FT 14480 14520
FT 14520 14560
FT 14560 14600
FT 14600 14640
FT 14640 14680
FT 14680 14720
FT 14720 14760
FT 14760 14800
FT 14800 14840
FT 14840 14880
FT 14880 14920
FT 14920 14960
FT 14960 15000
FT 15000 15040
FT 15040 15080
FT 15080 15120
FT 15120 15160
FT 15160 15200
FT 15200 15240
FT 15240 15280
FT 15280 15320
FT 15320 15360
FT 15360 15400
FT 15400 15440
FT 15440 15480
FT 15480 15520
FT 15520 15560
FT 15560 15600
FT 15600 15640
FT 15640 15680
FT 15680 15720
FT 15720 15760
FT 15760 15800
FT 15800 15840
FT 15840 15880
FT 15880 15920
FT 15920 15960
FT 15960 16000
FT 16000 16040
FT 16040 16080
FT 16080 16120
FT 16120 16160
FT 16160 16200
FT 16200 16240
FT 16240 16280
FT 16280 16320
FT 16320 16360
FT 16360 16400
FT 16400 16440
FT 16440 16480
FT 16480 16520
FT 16520 16560
FT 16560 16600
FT 16600 16640
FT 16640 16680
FT 16680 16720
FT 16720 16760
FT 16760 16800
FT 16800 16840
FT 16840 16880
FT 16880 16920
FT 16920 16960
FT 16960 17000
FT 17000 17040
FT 17040 17080
FT 17080 17120
FT 17120 17160
FT 17160 17200
FT 17200 17240
FT 17240 17280
FT 17280 17320
FT 17320 17360
FT 17360 17400
FT 17400 17440
FT 17440 17480
FT 17480 17520
FT 17520 17560
FT 17560 17600
FT 17600 17640
FT 17640 17680
FT 17680 17720
FT 17720 17760
FT 17760 17800
FT 17800 17840
FT 17840 17880
FT 17880 17920
FT 17920 17960
FT 17960 18000
FT 18000 18040
FT 18040 18080
FT 18080 18120
FT 18120 18160
FT 18160 18200
FT 18200 18240
FT 18240 18280
FT 18280 18320
FT 18320 18360
FT 18360 18400
FT 18400 18440
FT 18440 18480
FT 18480 18520
FT 18520 18560
FT 18560 18600
FT 18600 18640
FT 18640 18680
FT 18680 18720
FT 18720 18760
FT 18760 18800
FT 18800 18840
FT 18840 18880
FT 18880 18920
FT 18920 18960
FT 18960 19000
FT 19000 19040
FT 19040 19080
FT 19080 19120
FT 19120 19160
FT 19160 19200
FT 19200 19240
FT 19240 19280
FT 19280 19320
FT 19320 19360
FT 19360 19400
FT 19400 19440
FT 19440 19480
FT 19480 19520
FT 19520 19560
FT 19560 19600
FT 19600 19640
FT 19640 19680
FT 19680 19720
FT 19720 19760
FT 19760 19800
FT 19800 19840
FT 19840 19880
FT 19880 19920
FT 19920 19960
FT 19960 20000
FT 20000 20040
FT 20040 20080
FT 20080 20120
FT 20120 20160
FT 20160 20200
FT 20200 20240
FT 20240 20280
FT 20280 20320
FT 20320 20360
FT 20360 20400
FT 20400 20440
FT 20440 20480
FT 20480 20520
FT 20520 20560
FT 20560 20600
FT 20600 20640
FT 20640 20680
FT 20680 20720
FT 20720 20760
FT 20760 20800
FT 20800 20840
FT 20840 20880
FT 20880 20920
FT 20920 20960
FT 20960 21000
FT 21000 21040
FT 21040 21080
FT 21080 21120
FT 21120 21160
FT 21160 21200
FT 21200 21240
FT 21240 21280
FT 21280 21320
FT 21320 21360
FT 21360 21400
FT 21400 21440
FT 21440 21480
FT 21480 21520
FT 21520 21560
FT 21560 21600
FT 21600 21640
FT 21640 21680
FT 21680 21720
FT 21720 21760
FT 21760 21800
FT 21800 21840
FT 21840 21880
FT 21880 21920
FT 21920 21960
FT 21960 22000
FT 22000 22040
FT 22040 22080
FT 22080 22120
FT 22120 22160
FT 22160 22200
FT 22200 22240
FT 22240 22280
FT 22280 22320
FT 22320 22360
FT 22360 22400
FT 22400 22440
FT 22440 22480
FT 22480 22520
FT 22520 22560
FT 22560 22600
FT 22600 22640
FT 22640 22680
FT 22680 22720
FT 22720 22760
FT 22760 22800
FT 22800 22840
FT 22840 22880
FT 22880 22920
FT 22920 22960
FT 22960 23000
FT 23000 23040
FT 23040 23080
FT 23080 23120
FT 23120 23160
FT 23160 23200
FT 23200 23240
FT 23240 23280
FT 23280 23320
FT 23320 23360
FT 23360 23400
FT 23400 23440
FT 23440 23480
FT 23480 23520
FT 23520 23560
FT 23560 23600
FT 23600 23640
FT 23640 23680
FT 23680 23720
FT 23720 23760
FT 23760 23800
FT 23800 23840
FT 23840 23880
FT 23880 23920
FT 23920 23960
FT 23960 24000
FT 24000 24040
FT 24040 24080
FT 24080 24120
FT 24120 24160
FT 24160 24200
FT 24200 24240
FT 24240 24280
FT 24280 24320
FT 24320 24360
FT 24360 24400
FT 24400 24440
FT 24440 24480
FT 24480 24520
FT 24520 24560
FT 24560 24600
FT 24600 24640
FT 24640 24680
FT 24680 24720
FT 24720 24760
FT 24760 24800
FT 24800 24840
FT 24840 24880
FT 24880 24920
FT 24920 24960
FT 24960 25000
FT 25000 25040
FT 25040 25080
FT 25080 25120
FT 25120 25160
FT 25160 25200
FT 25200 25240
FT 25240 25280
FT 25280 25320
FT 25320 25360
FT 25360 25400
FT 25400 25440
FT 25440 25480
FT 25480 25520
FT 25520 25560
FT 25560 25600
FT 25600 25640
FT 25640 25680
FT 25680 25720
FT 25720 25760
FT 25760 25800
FT 25800 25840
FT 25840 25880
FT 25880 25920
FT 25920 25960
FT 25960 26000
FT 26000 26040
FT 26040 26080
FT 26080 26120
FT 26120 26160
FT 26160 26200
FT 26200 26240
FT 26240 26280
FT 26280 26320
FT 26320 26360
FT 26360 26400
FT 26400 26440
FT 26440 26480
FT 26480 26520
FT 26520 26560
FT 26560 26600
FT 26600 26640
FT 26640 26680
FT 26680 26720
FT 26720 26760
FT 26760 26800
FT 26800 26840
FT 26840 26880
FT 26880 26920
FT 26920 26960
FT 26960 27000
FT 27000 27040
FT 27040 27080
FT 27080 27120
FT 27120 27160
FT 27160 27200
FT 27200 27240
FT 27240 27280
FT 27280 27320
FT 27320 27360
FT 27360 27400
FT 27400 27440
FT 27440 27480
FT 27480 27520
FT 27520 27560
FT 27560 27600
FT 27600 27640
FT 27640 27680
FT 27680 27720
FT 27720 27760
FT 27760 27800
FT 27800 27840
FT 27840 27880
FT 27880 27920
FT 27920 27960
FT 27960 28000
FT 28000 28040
FT 28040 28080
FT 28080 28120
FT 28120 28160
FT 28160 28200
FT 28200 28240
FT 28240 28280
FT 28280 28320
FT 28320 28360
FT 28360 28400
FT 28400 28440
FT 28440 28480
FT 28480 28520
FT 28520 28560
FT 28560 28600
FT 28600 28640
FT 28640 28680
FT 28680 28720
FT 28720 28760
FT 28760 28800
FT 28800 28840
FT 28840 28880
FT 28880 28920
FT 28920 28960
FT 28960 29000
FT 29000 29040
FT 29040 29080
FT 29080 29120
FT 29120 29160
FT 29160 29200
FT 29200 29240
FT 29240 29280
FT 29280 29320
FT 29320 29360
FT 29360 29400
FT 29400 29440
FT 29440 29480
FT 29480 29520
FT 29520 29560
FT 29560 29600
FT 29600 29640
FT 29640 29680
FT 29680 29720
FT 29720 29760
FT 29760 29800
FT 29800 29840
FT 29840 29880
FT 29880 29920
FT 29920 29960
FT 29960 30000
FT 30000 30040
FT 30040 30080
FT 30080 30120
FT 30120 30160
FT 30160 30200
FT 30200 30240
FT 30240 30280
FT 30280 30320
FT 30320 30360
FT 30360 30400
FT 30400 30440
FT 30440 30480
FT 30480 30520
FT 30520 30560
FT 30560 30600
FT 30600 30640
FT 30640 30680
FT 30680 30720
FT 30720 30760
FT 30760 30800
FT 30800 30840
FT 30840 30880
FT 30880 30920
FT 30920 30960
FT 30960 31000
FT 31000 31040
FT 31040 31080
FT 31080 3112
```

```

FT DISULFID 115 131 BY SIMILARITY.
FT DISULFID 133 142 BY SIMILARITY.
FT DISULFID 148 159 BY SIMILARITY.
FT DISULFID 153 168 BY SIMILARITY.
FT DISULFID 170 179 BY SIMILARITY.
FT DISULFID 186 198 BY SIMILARITY.
FT DISULFID 192 207 BY SIMILARITY.
FT DISULFID 209 218 BY SIMILARITY.
FT DISULFID 225 236 BY SIMILARITY.
FT DISULFID 230 246 BY SIMILARITY.
FT DISULFID 248 257 BY SIMILARITY.

Query March 69.4%; Score 34; DB 1; Length 2471;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 VDNRECI 9
Db 1627 IDNRQCV 1633

RESULT 14
NTC2_RAT STANDARD; PRT; 2471 AA.
AC Q90M30;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neurogenic locus notch homolog protein 2 precursor (Notch 2).
GN NOTCH2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93202015; PubMed=1295745;
RA Weinmaster G., Roberts V.J., Lemke G.;
RT "Notch2: a second mammalian Notch gene.";
RL Development 116:931-941(1992).
[2]
RP TISSUE SPECIFICITY.
RX MEDLINE=21331789; PubMed=11438922;
RA Iryin D.K., Zurcher S.D., Nguyen T., Weinmaster G., Kornblum H.I.;
RT "Expression patterns of Notch1, Notch2, and Notch3 suggest multiple
functional roles for the Notch-DSL signaling system during brain
development.";
RL J. Comp. Neurol. 436:167-181(2001).
CC -1- FUNCTION: Functions as a receptor for membrane-bound ligands
Jagged1, Jagged2 and Deltal to regulate cell-fate determination.
Upon ligand activation through the released intracellular
domain (NICD) it forms a transcriptional activator complex with
RBP-J kappa and activates genes of the enhancer of split locus.
Affects the implementation of differentiation, proliferation and
apoptotic programs. May play an essential role in postimplantation
development, probably in some aspect of cell specification and/or
differentiation (By similarity).
CC -1- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
terminal fragment N(IEC) which are probably linked by disulfide
bonds (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Following
proteolytical processing NICD is translocated to the nucleus.
CC -1- TISSUE SPECIFICITY: Highly expressed in the spleen and choroid
plexus in the brain. Expressed in postnatal central nervous system
(CNS) germinal zones and, in early postnatal life, within numerous
cells throughout the CNS. It is more highly localized to
ventricular germinal zones. Also found in the heart, liver and
kidney.
CC -1- DEVELOPMENTAL STAGE: Expressed in the brain during E14 and E17.
CC -1- PTM: Synthesized in the endoplasmic reticulum as an inactive form
which is proteolytically cleaved by a furin-like convertase in the
trans-Golgi network before it reaches the plasma membrane to yield

```

```

CC an active, ligand-accessible form. Cleavage results in a C-
CC terminal fragment N(TM) and a N-terminal fragment N(IEC). Following
CC ligand binding, it is cleaved by TMP-alpha converting enzyme
CC (TACE) to yield a membrane-associated intermediate fragment called
CC notch extracellular truncation (NEXT). This fragment is then
CC cleaved by presenilin dependent gamma-secretase to release a
CC notch-derived peptide containing the intracellular domain (NICD)
CC from the membrane (By similarity).
CC -1- PTM: Phosphorylated (By similarity).
CC -1- SIMILARITY: Belongs to the NOTCH family.
CC -1- SIMILARITY: Contains 35 EGF-like domains.
CC -1- SIMILARITY: Contains 2 Lin/Notch repeats.
CC -1- SIMILARITY: Contains 6 ANK repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL: M93661; AAK13558.1; -.
CC PIR: A49128; A49128.
CC HSSP: P00743; 1CCF.
CC InterPro: IPR002110; ANK.
CC InterPro: IPR000152; Asx_hydroxyl_5.
CC InterPro: IPR000742; EGF_2.
CC InterPro: IPR001881; EGF_Ca.
CC InterPro: IPR001438; EGF_11.
CC InterPro: IPR006209; EGF_like.
CC InterPro: IPR002049; laminin_EGF.
CC InterPro: IPR008297; Notch.
CC InterPro: IPR008000; Notch_dom.
CC Pfam: PF00023; ank; 6.
CC Pfam: PF00008; EGF; 35.
CC Pfam: PF00066; notch; 2.
CC PIRSF: PIRSF02279; Notch; 1.
CC PRINTS: PR00010; EGFBLD.
CC PRINTS: PR00011; EGF_LAMININ.
CC SMART: SM01452; NOTCH.
CC SMART: SM00248; ANK; 6.
CC SMART: SM00179; EGF_CA; 24.
CC SMART: SM00004; NL_2.
CC PROSITE: PS50297; ANK_REPEAT; 1.
CC PROSITE: PS50088; ANK_REPEAT; 4.
CC PROSITE: PS00010; ASX_HYDROXYL; 22.
CC PROSITE: PS00022; EGF_1; 34.
CC PROSITE: PS01186; EGF_2; 26.
CC PROSITE: PS00026; EGF_3; 35.
CC PROSITE: PS01187; EGF_CA; 22.
CC Receptor: Transcription regulation; Activator; Differentiation;
CC developmental protein; Repeat; ANK repeat; EGF-like domain;
CC Transmembrane; Glycoprotein; Signal; Phosphorylation.
CC SIGNAL 1 25
CC CHAIN 26 2471
CC CHAIN 166 2471
CC CHAIN 1697 2471
CC DOMAIN 26 1677
CC TRANSFEM 1678 1698
CC DOMAIN 1699 2471
CC DOMAIN 26 63
CC DOMAIN 64 102
CC DOMAIN 105 143
CC DOMAIN 144 180
CC DOMAIN 182 219
CC DOMAIN 221 258
CC DOMAIN 260 296
CC DOMAIN 298 336
CC DOMAIN 338 374
CC DOMAIN 375 413

```

```

FT DOMAIN 415 454 EGF-LIKE 11, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 456 492 EGF-LIKE 12, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 494 530 EGF-LIKE 13, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 532 568 EGF-LIKE 14, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 570 605 EGF-LIKE 15, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 607 643 EGF-LIKE 16, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 645 680 EGF-LIKE 17, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 682 718 EGF-LIKE 18, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 720 755 EGF-LIKE 19, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 757 793 EGF-LIKE 20, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 795 831 EGF-LIKE 21, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 833 871 EGF-LIKE 22, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 873 909 EGF-LIKE 23, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 911 947 EGF-LIKE 24, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 949 985 EGF-LIKE 25, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 987 1023 EGF-LIKE 26, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1025 1061 EGF-LIKE 27, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1063 1099 EGF-LIKE 28, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1101 1147 EGF-LIKE 29, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1149 1185 EGF-LIKE 30, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1187 1223 EGF-LIKE 31, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1225 1262 EGF-LIKE 32, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1264 1302 EGF-LIKE 33, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1304 1343 EGF-LIKE 34, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1374 1412 EGF-LIKE 35, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1445 1484 EGF-LIKE 36, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1494 1535 EGF-LIKE 37, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1597 1648 EGF-LIKE 38, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1694 1745 EGF-LIKE 39, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1794 1845 EGF-LIKE 40, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1894 1945 EGF-LIKE 41, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 1994 2045 EGF-LIKE 42, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 2094 2145 EGF-LIKE 43, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 2194 2245 EGF-LIKE 44, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 2294 2345 EGF-LIKE 45, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 2394 2445 EGF-LIKE 46, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 2494 2545 EGF-LIKE 47, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 2594 2645 EGF-LIKE 48, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 2694 2745 EGF-LIKE 49, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 2794 2845 EGF-LIKE 50, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 2894 2945 EGF-LIKE 51, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 2994 3045 EGF-LIKE 52, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 3094 3145 EGF-LIKE 53, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 3194 3245 EGF-LIKE 54, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 3294 3345 EGF-LIKE 55, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 3394 3445 EGF-LIKE 56, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 3494 3545 EGF-LIKE 57, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 3594 3645 EGF-LIKE 58, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 3694 3745 EGF-LIKE 59, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 3794 3845 EGF-LIKE 60, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 3894 3945 EGF-LIKE 61, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 3994 4045 EGF-LIKE 62, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 4094 4145 EGF-LIKE 63, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 4194 4245 EGF-LIKE 64, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 4294 4345 EGF-LIKE 65, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 4394 4445 EGF-LIKE 66, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 4494 4545 EGF-LIKE 67, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 4594 4645 EGF-LIKE 68, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 4694 4745 EGF-LIKE 69, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 4794 4845 EGF-LIKE 70, CALCITUM-BINDING (POTENTIAL).
FT DOMAIN 4894 4945 EGF-LIKE 71, CALCITUM-BINDING (POTENTIAL).

```

```

FT DISULFID 498 509 BY SIMILARITY.
FT DISULFID 503 518 BY SIMILARITY.
FT DISULFID 520 529 BY SIMILARITY.
FT DISULFID 536 547 BY SIMILARITY.

Query Match
Best Local Similarity 57.1%; Score 34; DB 1; Length 2471;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNRECT 9
DB 1627 IDNRQCV 1633

RESULT 15
ID NTCL_MOUSE STANDARD; PRT: 2531 AA.
AC Q01705; Q06007; Q61905; Q99JC2; Q9QW58; Q9RQX7;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic locus notch homolog protein 1 precursor (Notch 1) (Notch A)
DE (mt14) (p300).
GN NOTCH1 OR NOTCH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10990;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Embryo;
RA MEDLINE=93194170; PubMed=8449489;
RA Franco del Amo F., Gendron-Maguire M., Swiatek P.J., Jenkins N.A.,
RA Copeland N.G., Gridley T.;
RT "Cloning, analysis, and chromosomal localization of Notch-1, a mouse
RT homolog of Drosophila Notch."
RT Genomics 15:259-264(1993).
RN [2]
RP SEQUENCE OF 711-1899 FROM N.A. (ISOFORM 2), AND DEVELOPMENTAL STAGE.
RC STRAIN=CD-1; TISSUE=Embryo;
RA MEDLINE=93050801; PubMed=1426644;
RA Reaume A.G., Conlon R.A., Ziringibi R., Yamaguchi T.P., Rossant J.;
RA "Expression analysis of a Notch homologue in the mouse embryo."
RA Dev. Biol. 154:377-387(1992).
RN [3]
RP SEQUENCE OF 1551-1647 FROM N.A. (ISOFORM 1), AND DEVELOPMENTAL STAGE.
RC TISSUE=Embryo;
RA MEDLINE=93048835; PubMed=1425352;
RA Franco del Amo F., Smith D.E., Swiatek P.J., Gendron-Maguire M.,
RA Greenspan R.J., McMahon A.P., Gridley T.;
RT "Expression pattern of Notch, a mouse homolog of Drosophila Notch,
RT suggests an important role in early postimplantation mouse
RT development."
RT Development 115:737-744(1992).
RN [4]
RP SEQUENCE OF 1161-1547 FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Embryo;
RA MEDLINE=931178563; PubMed=8440332;
RA Lardelli M., Lendahl U.;
RT "Notch A and Notch B-two mouse Notch homologues coexpressed in a
RT wide variety of tissues."
RT Exp. Cell Res. 204:364-372(1993).
RN [5]
RP SEQUENCE OF 1659-1673 FROM N.A.
RC MEDLINE=99364499; PubMed=10437788;
RA Lee J.S., Ishimoto A., Yanagawa S.I.;
RT "Murine leukemia provirus-mediated activation of the Notch1 gene leads
RT to induction of HES-1 in a mouse T lymphoma cell line, DJ-3."
RA FEBS Lett. 455:276-280(1999).
RN [6]
RP SEQUENCE OF 1950-2201 FROM N.A.
RC MEDLINE=98029496; PubMed=9384671;
RA Meserle M., Folio M., Nehls M., Eggert H., Boehm T.;

```

RT "Dynamic changes in gene expression during in vitro differentiation of
RT mouse embryonic stem cells.";
RT Cytokines Cell. Mol. Ther. 1:139-143(1995).
[7]
RP SEQUENCE OF 1655-1659, CLEAVAGE BY FURIN-LIKE CONVERTASE, AND
RP MUTAGENESIS OF 1651-ARG--ARG-1654.
RX MEDLINE=98318619; PubMed=9653148;
RA Logeat F., Bessia C., Brou C., Lebail O., Jarriault S., Seidah N.G.,
RA Israel A.;
RT "The Notch1 receptor is cleaved constitutively by a furin-like
RT convertase.";
RT Proc. Natl. Acad. Sci. U.S.A. 95:8108-8112(1998).
[8]
RP PARTIAL SEQUENCE, AND POST-TRANSLATIONAL PROCESSING.
RX MEDLINE=21523956; PubMed=11518718;
RA Saxena M.T., Schroeter E.H., Mumm J.S., Kopan R.;
RT "Murine notch homologs (NL-4) undergo presenilin-dependent
RT proteolysis.";
RL J. Biol. Chem. 276:40268-40273(2001).
[9]
RP POST-TRANSLATIONAL PROCESSING.
RX MEDLINE=21374376; PubMed=11459941;
RA Mizutani T., Taniguchi Y., Aoki T., Hashimoto N., Honjo T.;
RT "Conservation of the biochemical mechanisms of signal transduction
RT among mammalian Notch family members.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031(2001).
[10]
RP INTERACTION WITH DTX1 AND DTX2.
RX MEDLINE=21123790; PubMed=1126752;
RA Kishi N., Tang Z., Maeda Y., Hirai A., Mo R., Ito M., Suzuki S.,
RA Nakao K., Kinoshita T., Kadesch T., Hui C.-C., Aravanis-Tsakonas S.,
RA Okano H., Matsuno K.;
RT "Murine homologs of deltex define a novel gene family involved in
RT vertebrate Notch signaling and neurogenesis.";
RL Int. J. Dev. Neurosci. 19:21-35(2001).
-1- FUNCTION: Functions as a receptor for membrane-bound ligands
Jagged1, Jagged2 and Deltal to regulate cell-fate determination.
Upon ligand activation through the released notch intracellular
domain (NICD) it forms a transcriptional activator complex with
RBP-J kappa and activates genes of the enhancer of split locus.
Affects the implementation of differentiation, proliferation and
postimplantation programs (By similarity). May play an essential role in
specification and/or differentiation. May be involved in mesoderm
development, somite formation and neurogenesis. Involved in the
maturation of both CD4+ and CD8+ cells in the thymus.
-1- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
terminal fragment N(IEC) which are probably linked by disulfide
bonds. Interacts with DTX1 and DTX2.
-1- SUBCELLULAR LOCATION: Type I membrane protein. Following
proteolytical processing NICD is translocated to the nucleus.
-1- ALTERNATIVE PRODUCTS:
Event-Alternative splicing; Named isoforms-2;
Name-1;
IsoId=Q01705-1; Sequence=Displayed;
Name-2;
IsoId=Q01705-2; Sequence=VSP_001402, VSP_001403, VSP_001404;
Note=No experimental confirmation available;
-1- TISSUE SPECIFICITY: Highly expressed in the brain, lung and
thymus. Expressed at lower levels in the spleen, bone-marrow,
spinal cord, eyes, mammary gland, liver, intestine, skeletal
muscle, kidney and heart.
-1- DEVELOPMENTAL STAGE: First detected in the mesoderm at 7.5 dpc By
8.5 dpc highly expressed in presomitic mesoderm, mesenchyme and
endothelial cells, while much lower levels are seen in the
neuroepithelium. Between 9.5-10.5 dpc expressed at high levels in
the neuroepithelium. At 13.5 dpc expressed in the surface
ectoderm, eye and developing whisker follicles.
-1- PTM: Synthesized in the endoplasmic reticulum as an inactive form
which is proteolytically cleaved by a furin-like convertase to yield
an active, ligand-accessible form. Cleavage results in a C-
terminal fragment N(TM) and a N-terminal fragment N(IEC). Following

ligand binding, it is cleaved by TNF-alpha converting enzyme
(TACE) to yield a membrane-associated intermediate fragment called
notch extracellular truncation (NEXT). This fragment is then
cleaved by presenilin dependent gamma-secretase to release a
notch-derived peptide containing the intracellular domain (NICD)
from the membrane.
-1- PTM: Phosphorylated.
-1- SIMILARITY: Belongs to the NOTCH family.
-1- SIMILARITY: Contains 36 EGF-like domains.
-1- SIMILARITY: Contains 3 Lin/Notch repeats.
-1- SIMILARITY: Contains 5 ANK repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).

CC EMBL; Z11886; CAAT7941.1; -;
CC EMBL; I02613; AAK14898.1; -;
CC EMBL; X68278; CAA48339.1; -;
CC EMBL; AJ38029; CAB40733.1; -;
CC EMBL; X82562; CAA57909.1; -;
CC PIR; A46019; A46019.
CC PIR; B49175; B49175.
CC HSSP; P00740; 1EDM.
CC MGD; MGI:97363; Notch1.
CC GO; GO:0005887; C:integral to plasma membrane; IC.
CC GO; GO:0005515; P:protein binding; IPI.
CC GO; GO:0030154; P:cell differentiation; IMP.
CC GO; GO:0007386; P:compartment specification; IMP.
CC GO; GO:0007219; P:N signaling pathway; IC.
CC GO; GO:0045944; P:positive regulation of transcription from P. .; IDA.
CC InterPro; IPR002110; ANK.
CC InterPro; IPR000152; Asx hydroxyl_S.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR001438; EGF_II.
CC InterPro; IPR006209; EGF_III.
CC InterPro; IPR002049; Laminin_EGF.
CC InterPro; IPR008297; Notch.
CC InterPro; IPR000800; Notch_dom.
CC Pfam; PF00023; ank; 7.
CC Pfam; PF00066; notch; 3.
CC PIRSF; PIRSF02279; Notch; 1.
CC PRINTS; PR00010; EGFBLDOD.
CC PRINTS; PR00011; EGFFLAMININ.
CC SMART; SM00248; NOTCH.
CC SMART; SM00179; EGF_CA; 24.
CC SMART; SM00004; NLI_2.
CC PROSITE; PS50297; ANK_REPEAT; 1.
CC PROSITE; PS50088; ANK_REPEAT; 2.
CC PROSITE; PS00010; ASX_HYDROXYL; 22.
CC PROSITE; PS00022; EGF_1; 34.
CC PROSITE; PS01186; EGF_2; 27.
CC PROSITE; PS50026; EGF_3; 36.
CC PROSITE; PS01187; EGF_CA; 21.
CC Receptor; Transcription regulation; Activator; Differentiation;
CC Transmembrane; Glycoprotein; Signal; Phosphorylation;
CC Alternative splicing.
CC SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 2531 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1.
FT CHAIN 1711 2531 NOTCH EXTRACELLULAR TRUNCATION.
FT CHAIN 1744 2531 NOTCH INTRACELLULAR DOMAIN.
FT DOMAIN 19 1725 EXTRACELLULAR (POTENTIAL).

Query Match 69.4%; Score 34; DB 1; Length 2531;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 3 VNRECT 9
 Db 1670 INRQCV 1676

RESULT 16
 NCBI_RAT STANDARD; PRT; 2531 AA.
 ID NCBI_RAT
 AC Q07008;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neurogenic locus notch homolog protein 1 precursor (Notch 1).
 GN NOTCH1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Schwann cell;
 RX MEDLINE=92111383; PubMed=1764995;
 RA Weinmaster G., Roberts V.J., Lemke G.;
 RT "A homolog of Drosophila Notch expressed during mammalian
 development.";
 RL development 113:199-205 (1991).
 RN (2)
 RP REVISIONS TO 1652-1653.
 RA Weinmaster G.;
 RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP FUNCTION.
 RX MEDLINE=21094508; PubMed=1182080;
 RA Tanigaki K., Nogaki F., Takahashi J., Tashiro K., Kurooka H.,
 RA Honjo T.;
 RT "Notch1 and Notch3 instructively restrict bFGF-responsive multipotent
 neural progenitor cells to an astroglial fate.";
 RL Neuron 29:45-55 (2001).
 RN (4)
 RP TISSUE SPECIFICITY.
 RX MEDLINE=93202015; PubMed=1295745;
 RA Weinmaster G., Roberts V.J., Lemke G.;
 RT "Notch2: a second mammalian Notch gene.";
 RL development 116:931-941 (1992).
 RN (5)
 RP TISSUE SPECIFICITY.
 RX MEDLINE=21331789; PubMed=11438922;
 RA Irvin D.K., Zurcher S.D., Nguyen T., Weinmaster G., Kornblum H.I.;
 RT "Expression patterns of Notch1, Notch2, and Notch3 suggest multiple
 functional roles for the Notch-DSL signaling system during brain
 development.";
 RL J. Comp. Neurol. 436:167-181 (2001).
 CC -i- FUNCTION: Functions as a receptor for membrane-bound ligands
 Jagged1, Jagged2 and Delta to regulate cell-fate determination.
 Upon ligand activation through the released notch intracellular
 domain (NICD) it forms a transcriptional activator complex with
 RBP-J kappa and activates genes of the enhancer of split locus.
 Affects the implementation of differentiation, proliferation and
 apoptotic programs (By similarity). Acts instructively to control
 the cell fate determination of CNS multipotent progenitor cells,
 resulting in astroglial induction and neuron/oligodendrocyte
 suppression.
 CC -i- SUBUNIT: Heterodimer of a C-terminal fragment N(TW) and a N-
 terminal fragment N(EC) which are probably linked by disulfide
 bonds (By similarity).
 CC -i- SUBCELLULAR LOCATION: Type I membrane protein. Following
 proteolytical processing NICD is translocated to the nucleus (By
 similarity).
 CC -i- TISSUE SPECIFICITY: Expressed in the brain, kidney and spleen.
 CC Expressed in postnatal central nervous system (CNS) germinal zones
 and, in early postnatal life, within numerous cells throughout the

CC CNS. Found in both subventricular and ventricular germinal zones.
 CC -i- DEVELOPMENTAL STAGE: In the embryo, highest levels occur between
 CC days 12 and 14 and decrease rapidly to much lower levels in the
 CC adult.
 CC -i- PTM: Synthesized in the endoplasmic reticulum as an inactive form
 CC which is proteolytically cleaved by a furin-like convertase in the
 CC trans-golgi network before it reaches the plasma membrane to yield
 CC an active, ligand-accessible form. Cleavage results in a C-
 CC terminal fragment N(TW) and a N-terminal fragment N(EC). Following
 CC ligand binding, it is cleaved by TNF-alpha converting enzyme
 CC (TACE) to yield a membrane-associated intermediate fragment called
 CC notch extracellular truncation (NEXT). This fragment is then
 CC cleaved by presenilin dependent gamma-secretase to release a
 CC notch-derived peptide containing the intracellular domain (NICD)
 CC from the membrane (By similarity).
 CC -i- PTM: Phosphorylated (By similarity).
 CC -i- SIMILARITY: Belongs to the NOTCH family.
 CC -i- SIMILARITY: Contains 36 EGF-like domain.
 CC -i- SIMILARITY: Contains 3 Lin/Notch repeats.
 CC -i- SIMILARITY: Contains 5 ANK repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.iesb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X57405; CAA40667.1; -.
 CC HSSP: P00740; IEDM.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000152; Aex_hydroxyl_S.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001438; EGF_II.
 DR InterPro: IPR006209; EGF_like.
 DR InterPro: IPR002049; Laminin_EGF.
 DR InterPro: IPR008297; Notch.
 DR InterPro: IPR000800; Notch_dom.
 DR Pfam: PF00023; ank; 6.
 DR Pfam: PF00008; EGF; 35.
 DR Pfam: PF00066; notch; 3.
 DR PIRSF: PIRSF002279; Notch; 1.
 DR PRINTS: PR00010; EGFLOOD.
 DR PRINTS: PR00011; EGFPLAMIN.
 DR PRINTS: PR01452; NOTCH.
 DR SMART: SM00248; ANK; 6.
 DR SMART: SM00179; EGF_CA; 25.
 DR SMART: SM00004; NL; 2.
 DR PROSITE: PS50297; ANK_REPEAT; 1.
 DR PROSITE: PS00010; ASK_HYDROXYL; 22.
 DR PROSITE: PS00022; EGF_1; 35.
 DR PROSITE: PS01186; EGF_2; 26.
 DR PROSITE: PS50026; EGF_3; 36.
 DR PROSITE: PS01187; EGF_CA; 21.
 KM Receptor; Transcription regulation; Activator; Differentiation;
 KM Developmental protein; Repeat; ANK repeat; EGF-like domain;
 KM Transmembrane; Glycoprotein; Signal; Phosphorylation.
 FT SIGNAL 1 18
 FT CHAIN 19 2531
 FT CHAIN 1711 2531
 FT CHAIN 1744 2531
 FT CHAIN 19 1723
 FT TRANSMEM 1724 1746
 FT DOMAIN 1747 2531
 FT DOMAIN 20 58
 FT DOMAIN 59 99
 FT DOMAIN 102 139
 FT DOMAIN 140 176
 FT DOMAIN 140 176

FT	DOMAIN	178	216	EGF-LIKE 5,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	218	255	EGF-LIKE 6,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	257	293	EGF-LIKE 7,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	295	333	EGF-LIKE 8,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	335	371	EGF-LIKE 9,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	372	410	EGF-LIKE 10,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	412	450	EGF-LIKE 11,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	452	488	EGF-LIKE 12,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	490	526	EGF-LIKE 13,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	528	564	EGF-LIKE 14,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	566	601	EGF-LIKE 15,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	603	639	EGF-LIKE 16,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	641	676	EGF-LIKE 17,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	678	714	EGF-LIKE 18,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	716	751	EGF-LIKE 19,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	753	789	EGF-LIKE 20,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	791	827	EGF-LIKE 21,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	829	867	EGF-LIKE 22,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	869	905	EGF-LIKE 23,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	907	943	EGF-LIKE 24,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	945	981	EGF-LIKE 25,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	983	1019	EGF-LIKE 26,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1021	1057	EGF-LIKE 27,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1059	1095	EGF-LIKE 28,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1097	1143	EGF-LIKE 29,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1145	1181	EGF-LIKE 30,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1183	1219	EGF-LIKE 31,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1221	1265	EGF-LIKE 32,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1267	1305	EGF-LIKE 33,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1307	1346	EGF-LIKE 34,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1348	1384	EGF-LIKE 35,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1387	1426	EGF-LIKE 36,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1445	1480	LIN/NOTCH 1,	LIN/NOTCH 1.
FT	DOMAIN	1481	1522	LIN/NOTCH 2,	LIN/NOTCH 2.
FT	DOMAIN	1523	1562	LIN/NOTCH 3,	LIN/NOTCH 3.
FT	DOMAIN	1597	1946	ANK 1,	ANK 1.
FT	DOMAIN	1917	1946	ANK 2,	ANK 2.
FT	DOMAIN	1950	1980	ANK 3,	ANK 3.
FT	DOMAIN	1984	2013	ANK 4,	ANK 4.
FT	DOMAIN	2017	2046	ANK 5,	ANK 5.
FT	DOMAIN	2050	2079	POLY-ALA,	POLY-ALA.
FT	DOMAIN	1730	1733	POLY-GLU,	POLY-GLU.
FT	DOMAIN	1891	1894	POLY-PRO,	POLY-PRO.
FT	DOMAIN	2258	2261	CLEAVAGE BY (FURIN-LIKE PROTEASE) (BY	CLEAVAGE BY (FURIN-LIKE PROTEASE) (BY
FT	DOMAIN	2497	2500	SIMILARITY).	SIMILARITY).
FT	DOMAIN	1654	1655	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	24	37	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	31	46	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	48	57	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	63	74	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	68	87	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	89	98	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	106	117	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	111	127	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	129	138	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	144	155	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	149	164	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	166	175	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	182	195	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	189	204	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	206	215	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	222	233	BY SIMILARITY.	BY SIMILARITY.
FT	DOMAIN	227	243	BY SIMILARITY.	BY SIMILARITY.

Query Match 69.4%; Score 34; DB 1; Length 2531;
 Best Local Similarity 57.1%; Pred. 1.3e+02;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VDNRECT 9
 Db 1670 IDNRQCV 1676

RESULT 17
 ID NC1_HUMAN STANDARD; PRT: 2556 AA.
 AC P46531;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neurogenic locus notch homolog protein 1 precursor (Notch 1) (hN1)
 (translocation-associated notch protein TAN-1).
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Mann R.S., Blaumüller C.M., Zagouras P.;
 RT "Complete human notch 1 (hN1) cDNA sequence."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE OF 1-2444 FROM N.A.
 RX MEDLINE=91347367; PubMed=1831692;
 RA Ellisen L.W., Bird J., West D.C., Soreng A.L., Reynolds T.C.,
 RT Smith S.D., Sklar J.;
 RL "TAN-1, the human homolog of the Drosophila notch gene, is broken by
 chromosomal translocations in T lymphoblastic neoplasms."
 Cell 66:649-661(1991).
 RN (3)
 RP IDENTIFICATION OF LIGANDS.
 RX MEDLINE=99180765; PubMed=10079256;
 RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,
 RA Banks A., Leitman J., Ward D., Ish-Horowitz D., Aravanis-Teakonas S.;
 RT "Human ligands of the Notch receptor."
 RL Am. J. Pathol. 154:785-794(1999).
 RN (4)
 RP INTERACTION WITH DTX1.
 RX MEDLINE=98250176; PubMed=9590294;
 RA Matsuno K., Bastman D., Mitsiadis E., Quinn A.M., Carcangiu M.L.,
 RA Ordellich P., Kadesch T., Aravanis-Teakonas S.;
 RT "Human deltex is a conserved regulator of Notch signaling."
 Nat. Genet. 19:74-78(1998).
 RL
 CC -1- FUNCTION: Functions as a receptor for membrane-bound ligands
 tagged1, tagged2 and Deltal to regulate cell-fate determination.
 CC Upon ligand activation through the released notch intracellular
 CC domain (NICD) it forms a transcriptional activator complex with
 CC RBP-J kappa and activates genes of the enhancer of split locus.
 CC Affects the implementation of differentiation, proliferation and
 CC apoptotic programs. May be important for normal lymphocyte
 CC function. In altered form, may contribute to transformation or
 CC progression in some T-cell neoplasms. Involved in the maturation
 CC of both CD4+ and CD8+ cells in the thymus (By similarity).
 CC -1- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
 CC terminal fragment N(EC) which are probably linked by disulfide
 CC bonds (By similarity). Interacts with DTX1 and DTX2.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Following
 CC proteolytical processing NICD is translocated to the nucleus (By
 CC similarity).
 CC -1- TISSUE SPECIFICITY: In fetal tissues most abundant in spleen,
 CC brain stem and lung. Also present in most adult tissues where it
 CC is found mainly in lymphoid tissues.
 CC -1- PTM: Synthesized in the endoplasmic reticulum as an inactive form
 CC which is proteolytically cleaved by a furin-like convertase in the
 CC trans-Golgi network before it reaches the plasma membrane to yield
 CC an active, ligand-accessible form. Cleavage results in a C-
 CC terminal fragment N(TM) and a N-terminal fragment N(EC). Following
 CC ligand binding, it is cleaved by TMP-alpha converting enzyme
 CC (TACE) to yield a membrane-associated intermediate fragment called
 CC notch extracellular truncation (NEXT). This fragment is then
 CC cleaved by presenilin dependent gamma-secretase to release a
 CC notch-derived peptide containing the intracellular domain (NICD)
 CC from the membrane (By similarity).
 CC -1- PTM: Phosphorylated (By similarity).

CC -!- DISBASE: NOTCH1 truncation is associated with T-cell acute
 CC lymphoblastic leukemia.
 CC -!- SIMILARITY: Belongs to the NOTCH family.
 CC -!- SIMILARITY: Contains 36 EGF-like domains.
 CC -!- SIMILARITY: Contains 3 Lin/Notch repeats.
 CC -!- SIMILARITY: Contains 5 ANK repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF308602; AAG3848.1; -
 CC EMBL; M73980; AAA60614.1; -
 CC HSSP; P00740; 1EDM.
 CC Genew; HGNC:7881; NOTCH1.
 CC MIM; 190198; -
 CC GO; GO:0016021; C:integral to membrane; NAS.
 CC GO; GO:0003793; F:defense/immunity protein activity; NAS.
 CC GO; GO:0006955; P:immune response; NAS.
 CC InterPro; IPR002110; ANK.
 CC InterPro; IPR000152; Asx hydroxyl_S.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR001881; EGF_Ca.
 CC InterPro; IPR001438; EGF_II.
 CC InterPro; IPR006209; EGF-like.
 CC InterPro; IPR002049; laminin_EGF.
 CC InterPro; IPR008297; Notch.
 CC InterPro; IPR000800; Notch_dom.
 CC Pfam; PF00023; ank; 6.
 CC Pfam; PF00068; notch; 3.
 CC PIRSF; PIRSF002279; Notch; 1.
 CC PRINTS; PRO0010; EGFLCLOD.
 CC PRINTS; PRO0011; EGFLAMININ.
 CC PRINTS; PRO1452; NOTCH.
 CC SMART; SM00248; ANK; 6.
 CC SMART; SM00179; EGF_CA; 23.
 CC SMART; SM00004; NL; 3.
 CC PROSITE; PS50297; ANK_REPEAT; 1.
 CC PROSITE; PS50088; ANK_REPEAT; 4.
 CC PROSITE; PS00010; ASX_HYDROXYL; 20.
 CC PROSITE; PS00022; EGF_1; 34.
 CC PROSITE; PS01186; EGF_2; 26.
 CC PROSITE; PS50026; EGF_3; 36.
 CC PROSITE; PS01187; EGF_CA; 18.
 CC Receptor; Transcription regulation; Activator; Differentiation;
 CC developmental protein; Repeat; ANK repeat; EGF-like domain;
 CC Transmembrane; Glycoprotein; Signal; Phosphorylation.
 CC SIGNAL; 18
 CC CHAIN; 19 2556
 CC CHAIN; 1722 2556
 CC CHAIN; 1755 2556
 CC DOMAIN; 19 1736
 CC TRANSMEM; 1737 1757
 CC DOMAIN; 1758 2556
 CC DOMAIN; 20 58
 CC DOMAIN; 59 99
 CC DOMAIN; 102 139
 CC DOMAIN; 140 176
 CC DOMAIN; 178 216
 CC DOMAIN; 218 255
 CC DOMAIN; 257 293
 CC DOMAIN; 295 333
 CC DOMAIN; 335 371
 CC DOMAIN; 372 410
 CC DOMAIN; 412 450
 CC DOMAIN; 452 488

FT DOMAIN 490 526 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 528 564 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 566 601 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 603 639 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 641 676 EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 678 714 EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 716 751 EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 753 789 EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 791 827 EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 829 868 EGF-LIKE 22, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 870 906 EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 908 944 EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 946 982 EGF-LIKE 25, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1020 1058 EGF-LIKE 26, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1060 1096 EGF-LIKE 28, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1098 1144 EGF-LIKE 29, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1146 1182 EGF-LIKE 30, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1184 1220 EGF-LIKE 31, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1222 1266 EGF-LIKE 32, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1268 1306 EGF-LIKE 33, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1308 1347 EGF-LIKE 34, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1349 1385 EGF-LIKE 35, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1427 1461 EGF-LIKE 36, CALCIUM-BINDING (POTENTIAL).
 FT REPEAT 1446 1481 LIN/NOTCH 1.
 FT REPEAT 1482 1523 LIN/NOTCH 2.
 FT REPEAT 1524 1563 LIN/NOTCH 3.
 FT REPEAT 1565 1597 ANK 1.
 FT REPEAT 1597 1635 ANK 2.
 FT REPEAT 1635 1665 ANK 3.
 FT REPEAT 1665 1695 ANK 4.
 FT REPEAT 1695 1729 ANK 5.
 FT DOMAIN 1729 1765 POLY-VAL.
 FT DOMAIN 1765 1800 POLY-ARG.
 FT DOMAIN 1800 1832 POLY-PRO.
 FT DOMAIN 1832 1865 POLY-ALA.
 FT DOMAIN 1865 1905 POLY-GLU.
 FT DOMAIN 1905 1940 POLY-GLY.
 FT DOMAIN 1940 1975 POLY-GLN.
 FT DOMAIN 1975 2010 POLY-PRO.
 FT DOMAIN 2010 2045 POLY-SER.
 FT DOMAIN 2045 2080 CLEAVAGE (BY FURIN-LIKE PROTEASE) (BY
 FT SITE 1665 1666 SIMILARITY).
 FT DISULFID 24 37 BY SIMILARITY.
 FT DISULFID 31 46 BY SIMILARITY.
 FT DISULFID 48 57 BY SIMILARITY.
 FT DISULFID 63 74 BY SIMILARITY.
 FT DISULFID 89 98 BY SIMILARITY.
 FT DISULFID 106 117 BY SIMILARITY.
 FT DISULFID 111 127 BY SIMILARITY.
 FT DISULFID 129 138 BY SIMILARITY.
 FT DISULFID 144 155 BY SIMILARITY.
 FT DISULFID 149 164 BY SIMILARITY.
 FT DISULFID 166 175 BY SIMILARITY.
 FT DISULFID 182 195 BY SIMILARITY.
 Query Match Score 34; DB 1; Length 2556;
 Best Local Similarity 57.1%; Pred. No. 1.3e+02;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 3 VDNRECI 9
 Db 1681 IDNRQCV 1687
 RESULT 18
 ID KITH_MOUSE STANDARD; PRT; 233 AA.
 AC P04184;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Thymidine kinase, cytosolic (EC 2.7.1.21).
 GN TK1 OR TK-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=86310784; PubMed=3018504;
 RA Lin P.-F., Lieberman H.B., Yeh D.-B., Xu T., Zhao S.-Y., Ruddle F.H.;
 RT "Molecular cloning and structural analysis of murine thymidine kinase
 RT genomic and cDNA sequences.";
 RL Mol. Cell. Biol. 5:3149-3156(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92380505; PubMed=1511894;
 RA Gudas J.M., Fridovich-Keil J.L., Datta M.W., Bryan J., Pardee A.B.;
 RT "Characterization of the murine thymidine kinase-encoding gene and
 RT analysis of transcription start point heterogeneity.";
 RL Gene 118:205-216(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93211435; PubMed=7681542;
 RA Liechty M.C., Rauchfuss H.S., Lugo M.H., Hozier J.C.;
 RT "Sequence analysis of tk(-)-1 and tkb(-)-1 alleles in L5178Y tk+/-
 RT mouse-lymphoma cells and spontaneous tk-/-mutants.";
 RL Mutat. Res. 286:299-307(1993).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92359027; PubMed=1496920;
 RA Wintersberger E., Rotheneder H., Grabner M., Beck G., Seiser C.;
 RT "Regulation of thymidine kinase during growth, cell cycle and
 RT differentiation.";
 RL Adv. Enzyme Regul. 32:241-254(1992).
 RN [5]
 RP SEQUENCE OF 1-32 FROM N.A.
 RA TISSUE=Liver.
 RX MEDLINE=92107666; PubMed=1762910;
 RA Rotheneder H., Grabner M., Wintersberger E.;
 RT "Presence of regulatory sequences within intron 2 of the mouse
 RT thymidine kinase gene.";
 RL Nucleic Acids Res. 19:6805-6809(1991).
 RN [6]
 RP SEQUENCE OF 1-27 FROM N.A.
 RX MEDLINE=89218988; PubMed=3244356;
 RA Lieberman H.B., Lin P.F., Yeh D.B., Ruddle F.H.;
 RT "Transcriptional and posttranscriptional mechanisms regulate murine
 RT thymidine kinase gene expression in serum-stimulated cells.";
 RL Mol. Cell. Biol. 8:5280-5291(1988).
 CC -1- CATALYTIC ACTIVITY: ATP + thymidine = ADP + thymidine 5'-
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- PM: Phosphorylated on Ser-13 in mitosis (By similarity).
 CC -1- MISCELLANEOUS: TWO FORMS HAVE BEEN IDENTIFIED IN ANIMAL CELLS, ONE
 CC IN CYTOSOL & ONE IN MITOCHONDRIA. ACTIVITY OF THE CYTOSOLIC ENZYME
 CC IS HIGH IN PROLIFERATING CELLS & PEAKS DURING THE S PHASE OF THE
 CC CELL CYCLE; IT IS VERY LOW IN RESTING CELLS.
 CC -1- SIMILARITY: Belongs to the thymidine kinase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M11945; AAA40451.1; -;
 CC EMBL; M11945; AAA40451.1; ALT _INIT.
 CC EMBL; M68489; AAA40454.1; -;
 CC EMBL; S57244; AAB26001.1; -;
 CC EMBL; X60980; CAA43296.1; -;
 CC EMBL; M22581; AAA40450.1; -;

DR PIR; JCI252; KIMST.
 DR MGD; MGI:98763; TK1.
 DR InterPro; IPR001267; TK_cell.
 DR Pfam; PF00265; TK_1.
 DR PROSITE; PS00603; TK_CELLULAR_TYPE; 1.
 KW Transferase; Kinase; DNA synthesis; ATP-binding; Phosphorylation.
 FT NP BIND 26 33
 FT MOD RES 13 13
 FT VARIANT 88 88
 FT CONFLICT 88 88
 SQ SEQUENCE 233 AA; 25760 MW; 790946FEFCDDC72 CRC64;
 Query Match 67.34; Score 33; DB 1; Length 233;
 Best Local Similarity 55.64; Pred. No. 18;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AGVDNRECI 9
 DB 198 AGSDNKNCL 206
 RESULT 19
 ID EPH1 HUMAN STANDARD; PRT; 976 AA.
 AC P21709; O15405;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ephrin type-A receptor 1 precursor (EC 2.7.1.112) (Tyrosine-protein
 DE kinase receptor Eph).
 GN EPH1 OR EPH1 OR EPHT OR EPH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88070650; PubMed=2825356;
 RA Hirai H., Maru Y., Hagihara K., Nishida J., Takaku F.;
 RT "A novel putative tyrosine kinase receptor encoded by the eph gene.";
 RL Science 238:1717-1720(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99299440; PubMed=10369740;
 RA Owehalimpr D., Kelley M.J.;
 RT "Genomic structure of the EPH1 receptor tyrosine kinase gene.";
 RL Mol. Cell. Probes 13:169-173(1999).
 RN [3]
 RP SEQUENCE OF 286-976 FROM N.A.
 RC TISSUE=Placenta;
 RA Tuzi N.L.;
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS WITH
 CC A LOW AFFINITY TO EPHRIN-A1.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: OVEREXPRESSED IN SEVERAL CARCINOMAS.
 CC -1- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
 CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
 CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Ephrin
 CC receptor subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M18391; AAA36747.1; ALT_SEQ.
 CC EMBL; AF101171; AAD43440.1; -;


```

OX  NCBI_TaxID=9913;
RN  [1]
RP  SEQUENCE.
RC  TISSUE=Testis;
RA  MEDLINE=95262712; PubMed=7744071;
RT  Malpelli G., Stopini M., Zapponi M.C., Foll C., Berni R.;
RT  "Interactions with retinol and retinoids of bovine cellular retinol-
RT  binding protein.";
RL  Eur. J. Biochem. 229:486-493 (1995).
RN  [2]
RP  SEQUENCE OF 1-30.
RC  TISSUE=Retina;
RA  MEDLINE=82027750; PubMed=6269887;
RT  Crabb J.W., Saari J.C.;
RT  "N-terminal sequence homology among retinoid-binding proteins from
RT  bovine retina.";
RL  FEBS Lett. 130:15-18 (1981).
CC  -1- FUNCTION: Intracellular transport of retinol.
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -1- SIMILARITY: Belongs to the fatty-acid binding protein (FABP)
CC  family.
DR  PIR; S69360; S69360.
DR  HSSP; P02696; ICRB.
DR  InterPro; IPR000463; Fatty acid BP.
DR  InterPro; IPR000566; Lipocln_cytfabp.
DR  Pfam; PF00061; lipocalin; 1.
DR  PRINTS; PR00178; FATTYACIDBP.
DR  PROSITE; PS00214; FABP_1.
KW  Vitamin A; Retinol-binding; Transport.
FT  INIT MET
FT  CONFLICT
SQ  SEQUENCE 134 AA; 15699 MW; B20F84239879B7A CRC64;

Query Match
Best Local Similarity 50.0%; Score 32; DB 1; Length 134;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY  2 GVDNRECT 9
Db  76 GIDDRKCM 83

RESULT 22
RETI_HUMAN STANDARD; PRT; 134 AA.
ID RETI_HUMAN
AC P09455;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Retinol-binding protein I, cellular (Cellular retinol-binding protein)
DE (CRBP).
GN RBP1 OR CRBP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN  [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=85279409; PubMed=2992469;
RC Colantoni V., Cortese R., Nilsson M., Lundvall J., Baavik C.-O.,
RA Eriksson U., Peterson P.A., Sundelin J.;
RT "Cloning and sequencing of a full length cDNA corresponding to human
RT cellular retinol-binding protein.";
RL Biochem. Biophys. Res. Commun. 130:431-439 (1985).
RN  [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=88185326; PubMed=3356192;
RC Nilsson M.H.L., Spurr N.K., Lundvall J., Raak L., Peterson P.A.;
RT "Human cellular retinol-binding protein gene organization and
RT chromosomal location.";
RL Eur. J. Biochem. 173:35-44 (1988).
RN  [3]
RP SEQUENCE OF 1-120 FROM N.A.

```

```

RX  MEDLINE=91042593; PubMed=2856408;
RA  Wei L.N., Mertz J.R., Goodman D.S., Nguyen-Huu M.C.;
RT  "Cellular retinoic acid- and cellular retinol-binding proteins:
RT  complementary deoxyribonucleic acid cloning, chromosomal assignment,
RT  and tissue specific expression.";
RL  Mol. Endocrinol. 1:526-534 (1987).
RN  [4]
RP TISSUE SPECIFICITY.
RX  MEDLINE=21173623; PubMed=11274389;
RA  Foll C., Calderone V., Ottoneillo S., Bolchi A., Zanotti G.,
RA  Stopini M., Berni R.;
RT  "Identification, retinoid binding and X-ray analysis of a human
RT  retinol-binding protein.";
RL  Proc. Natl. Acad. Sci. U.S.A. 98:3710-3715 (2001).
CC  -1- FUNCTION: Intracellular transport of retinol.
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -1- TISSUE SPECIFICITY: DETECTED IN NEARLY ALL THE TISSUES WITH HIGHER
CC  EXPRESSION IN ADULT OVARY, PANCREAS, PITUITARY GLAND AND ADRENAL
CC  GLAND, AND FETAL LIVER.
CC  -1- SIMILARITY: Belongs to the fatty-acid binding protein (FABP)
CC  family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by, and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M11433; AAA60257.1; -.
DR  EMBL; X07437; CA30318.1; -.
DR  EMBL; X07438; CA30318.1; JOINED.
DR  EMBL; M36809; AAA35714.1; -.
DR  PIR; S00399; RJHUO.
DR  HSSP; P02696; ICRB.
DR  Genew; HGNC:9919; RBP1.
DR  MIM; 180260; -.
DR  GO; GO:0005501; F:retinoid binding; TAS.
DR  GO; GO:0006776; P:vitamin A metabolism; TAS.
DR  InterPro; IPR000463; Fatty acid BP.
DR  InterPro; IPR000566; Lipocln_cytfabp.
DR  Pfam; PF00061; lipocalin; 1.
DR  PRINTS; PR00178; FATTYACIDBP.
DR  PROSITE; PS00214; FABP_1.
KW  Vitamin A; Retinol-binding; Transport.
FT  INIT MET
FT  CONFLICT
SQ  SEQUENCE 134 AA; 15719 MW; A57E54573E969A6B CRC64;

Query Match
Best Local Similarity 50.0%; Score 32; DB 1; Length 134;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY  2 GVDNRECT 9
Db  76 GIDDRKCM 83

RESULT 23
RETI_MOUSE STANDARD; PRT; 134 AA.
ID RETI_MOUSE
AC Q00915;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Retinol-binding protein I, cellular (Cellular retinol-binding protein)
DE (CRBP) (mCRBP1).
GN RBP1 OR RBP-1 OR CRBP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN  [1]

```

RP SEQUENCE FROM N.A.
 RC TISSUE=Testis; PubMed=1648481; Rees J., Chambon P.;
 RA MEDLINE=91293098; PubMed=1648481; Rees J., Chambon P.;
 RA Smith W.C., Nakshatri H., Leroy P.,
 RT "A retinoic acid response element is present in the mouse cellular
 RT retinol binding protein I (mCRBP) promoter";
 RL EMBO J. 10:2223-2230(1991).
 CC -1- FUNCTION: Intracellular transport of retinol.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the fatty-acid binding protein (FABP)
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X60367; CAA42919.1; --
 DR PIR: S16355; S16355.
 DR HSSP: P02696; 1CRB.
 DR MGD: MGI:97876; RBP1.
 DR InterPro: IPR000463; Fatty acid BP.
 DR InterPro: IPR000566; Lipocin_cytfabp.
 DR Pfam: PF00061; Lipocalin_1.
 DR PRINTS: PR00178; FATVACIDBP.
 DR PROSITE: PS00214; FABP; 1.
 DR Vitamin A; Retinol-binding; Transport.
 DR INIT MET 0 BY SIMILARITY.
 SQ SEQUENCE 134 AA; 15715 MW; 372FE72DB84117E CRC64;
 Query Match 65.3%; Score 32; DB 1; Length 134;
 Best Local Similarity 50.0%; Pred. No. 16;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 GYDNRCCI 9
 DB 76 GIDDRKCM 83

RESULT 24
 REPT RAT STANDARD; PRT; 134 AA.
 ID REPT RAT
 AC P02696;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Retinol-binding protein I, cellular (cellular retinol-binding protein)
 DE (CRBP).
 GN RBP1 OR RBP-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OK NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87204107; PubMed=3472205;
 RA Sherman D.R., Lloyd R.S., Chytil F.;
 RT "Rat cellular retinol-binding protein: cDNA sequence and rapid
 RT retinol-dependent accumulation of mRNA";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:3209-3213(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87222311; PubMed=3584109;
 RA Levin M.S., Li E., Ong D.E., Gordon J.I.;
 RT "Comparison of the tissue-specific expression and developmental
 RT regulation of two closely linked rodent genes encoding cytosolic
 RT retinol-binding proteins";
 RL J. Biol. Chem. 262:7118-7124(1987).
 RN [3]
 RP SEQUENCE.

RC TISSUE=Liver;
 RX MEDLINE=85207645; PubMed=4039728;
 RA Sundelin U., Anundi H., Traegardh L., Eriksson U., Lind P., Ronne H.,
 RA Peterson P.A., Rask L.;
 RT "The primary structure of rat liver cellular retinol-binding
 RT protein";
 RL J. Biol. Chem. 260:6488-6493(1985).
 RN [4]
 RP SEQUENCE OF 1-50.
 RC TISSUE=Testis; PubMed=6541654;
 RX MEDLINE=85030476; PubMed=6541654;
 RA Eriksson U., Das K., Busch C., Nordlinder H., Rask L., Sundelin J.,
 RA Sallstrom J., Peterson P.A.;
 RT "Cellular retinol-binding protein. Quantitation and distribution";
 RL J. Biol. Chem. 259:13464-13470(1984).
 RN [5]
 RP SEQUENCE OF 1-34.
 RC TISSUE=Liver;
 RX MEDLINE=81254137; PubMed=6942701;
 RA Rask L., Anundi H., Boehme J., Eriksson U., Ronne H., Sege K.,
 RA Peterson P.A.;
 RT "Structural and functional studies of vitamin A-binding proteins";
 RL Ann. N.Y. Acad. Sci. 359:79-90(1981).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
 RX MEDLINE=93253782; PubMed=7683727;
 RA Cowan S.W., Newcomer M.E., Jones T.A.;
 RT "Crystallographic studies on a family of cellular lipophilic
 RT transport proteins. Refinement of P2 myelin protein and the structure
 RT determination and refinement of cellular retinol-binding protein in
 RT complex with all-trans-retinol";
 RL J. Mol. Biol. 230:1225-1246(1993).
 CC -1- FUNCTION: Intracellular transport of retinol.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the fatty-acid binding protein (FABP)
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M16459; AAA42021.1; --
 DR EMBL: M19257; AAA40962.1; --
 DR PIR: A29570; RJRTO.
 DR PDB: 1CRB; 20-DEC-94.
 DR PDB: 1UBH; 19-JUN-02.
 DR PDB: 1KGL; 03-JUL-02.
 DR InterPro: IPR000463; Fatty acid BP.
 DR InterPro: IPR000566; Lipocin_cytfabp.
 DR Pfam: PF00061; Lipocalin_1.
 DR PRINTS: PR00178; FATVACIDBP.
 DR PROSITE: PS00214; FABP; 1.
 DR Vitamin A; Retinol-binding; Transport; 3D-structure.
 KW INIT MET 0
 FT STRAND 6 14
 FT HELIX 16 21
 FT TURN 22 24
 FT TURN 27 35
 FT HELIX 39 45
 FT STRAND 46 47
 FT TURN 48 54
 FT STRAND 55 60
 FT TURN 66 67
 FT TURN 70 73
 FT STRAND 75 78
 FT TURN 80 89
 FT STRAND 91 91
 FT TURN 92 98
 FT TURN 102 103

[illegible]

MEDLINE=21848401; PubMed=11859360;
 Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 Sgouros J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,
 Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 Collins M., Connor A., Cronin A., Davis P., Felwell T., Fraser A.,
 Genies S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jaseels K.,
 James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 Mooney P., Mount S., Mungall K., Murphy L., Niblett D., Ocell C.,
 Oliver K., O'Neill S., Pearson S., Saunders D., Seeger K., Sharp S.,
 Rutherford K., Rutter S., Squares R., Squares S., Stevens K.,
 Skelton J., Simmonds M., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 Woodward J., Wolckart G., Aert R., Robben J., Grympey B.,
 Wellens I., Vanstreels B., Rieger M., Schaefer M., Mueller-Auer S.,
 Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 Eger P., Zimmermann W., Medler H., Mamburt R., Purnelle B.,
 Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forzburg S.L.,
 Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 "The genome sequence of *Schizosaccharomyces pombe*,"
 Nature 415:871-880(2002).
 [5]
 IDENTIFICATION AS 26S PROTEASOME SUBUNIT, AND FUNCTION.
 MEDLINE=98395111; PubMed=9727008;
 Penney M., Wilkinson C., Wallace M., Javerzat J.P., Ferrell K.,
 Seeger M., Dubiel W., McKay S., Alleghire R., Gordon C.;
 "The pad1 gene encodes a subunit of the 26 S proteasome in fission
 yeast.";
 J. Biol. Chem. 273:23938-23945(1998).
 -1- FUNCTION: Acts as a regulatory subunit of the 26 proteasome which
 is involved in the ATP-dependent degradation of ubiquitinated
 proteins.
 -1- FUNCTION: Transcription factor pap1 is controlled by the
 functional interaction between the positive regulator pad1 and
 negative regulator crml. Both these proteins are also essential
 for cell viability and for the maintenance of chromosome
 structure. Pad1 is also responsible for resistance to
 staurosporine and other drugs such as cycloheximide and caffeine.
 -1- SUBUNIT: The 26S proteasome is composed of a core protease, known
 as the 20S proteasome, capped at one or both ends by the 19S
 regulatory complex (RC). The RC is composed of at least 18
 different subunits in two subcomplexes, the base and the lid,
 which form the portions proximal and distal to the 20S proteolytic
 core, respectively (by similarity).
 -1- SIMILARITY: Contains 1 MPN (JAB/Mov34) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: D31731; BAA06529.1; -
 EMBL: D45047; BAA08087.1; -
 EMBL: D84656; BAA12708.1; -
 EMBL: Z98979; CAB11697.1; -
 DR PIR: T43293; T43293;
 DR PIR: T44427; T44427;
 DR GeneDB SPombe: SPAC3165.13; -
 DR GO: GO:0005837; C:26S proteasome; ISS.
 DR GO: GO:0005894; C:chromosome; ISS.
 DR GO: GO:0005829; C:cytosol; ISS.
 DR GO: GO:0005654; C:nucleoplasm; ISS.
 DR GO: GO:0006464; P:protein modification; ISS.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; ISS.

DR InterPro: IPR003639; Pept M67 Mov34.
 DR InterPro: IPR000555; Peptidase_M67.
 DR Pfam: PF01398; Mov34_1; 1.
 DR ProDom: PD363422; Mov34_1; 1.
 DR SMART: SM00232; JAB_MPN; 1.
 DR Proteasome: Antidiotic resistance; Cycloheximide resistance.
 FT CONFLICT 112 112 H -> N (IN REF. 2).
 SQ SEQUENCE 308 AA; 34572 MW; 57827C556B34E214 CRC64;

 Query Match 65.3%; Score 32; DB 1; Length 308;
 Best Local Similarity 71.4%; Pred. No. 38;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

 QY 3 VDNRECI 9
 Db 24 VDNSECV 30

 RESULT 27
 VIL HPV54 STANDARD; PRT; 497 AA.
 ID VIL HPV54
 AC P50819; Q81024;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Major capsid protein L1.
 GN L1.
 OS Human papillomavirus type 54.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCB1_TaxID=37113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Delius H.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 RT [2]
 RP SEQUENCE OF 310-460 FROM N.A.
 RX MEDLINE=95052821; PubMed=7963696;
 RA Bernard H.U., Chan S.Y., Manos M.M., Ong C.K., Villa L.L.,
 Delius H., Peyton C.L., Bauer H.M., Wheeler C.M.;
 "Identification and assessment of known and novel human
 papillomaviruses by polymerase chain reaction amplification,
 restriction fragment length polymorphisms, nucleotide sequence, and
 phylogenetic algorithms.";
 RT J. Infect. Dis. 170:1077-1085(1994).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL: U37488; AAA79193.1; -
 CC EMBL: U12501; AAA67245.1; -
 DR InterPro: IPR002210; PV_capsid_L1.
 DR InterPro: IPR008975; Viral_cap_coat.
 DR Pfam: PF00500; late_protein_L1; 1.
 DR PRINTS: PR000865; HPVCAPIIDL1.
 DR ProDom: PD000544; PV_capsid_L1; 1.
 KM Coat protein; late protein.
 FT CONFLICT 313 313 Q -> H (IN REF. 2).
 FT CONFLICT 323 323 V -> L (IN REF. 2).
 FT CONFLICT 368 368 T -> T (IN REF. 2).
 FT CONFLICT 375 375 I -> A (IN REF. 2).
 FT CONFLICT 439 439 N -> T (IN REF. 2).
 FT CONFLICT 456 456 F -> Y (IN REF. 2).
 SQ SEQUENCE 497 AA; 55840 MW; A6F7A015F8476A4B CRC64;

 Query Match 65.3%; Score 32; DB 1; Length 497;
 Best Local Similarity 85.7%; Pred. No. 62;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AGVDNRE 7
 DB 135 AGADNRE 141

RESULT 28

CHLB_EPHAL STANDARD; PRT; 509 AA.

AC P37846;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Light-independent protochlorophyllide reductase subunit B
 (EC 1.18.-.-) (LI-POR subunit B) (DPOK subunit B).
 GN CHLB.
 OS Eupheda altissima.
 OG Chloroplast.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Gnecophyta; Gnecopsida; Ephedrales; Ephedraceae;
 CC Ephedra.
 CC NCB1_TaxID=3391;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96426856; PubMed=8812302;
 RA Boivin R., Richard M., Beauseigle D., Bousquet J., Bellemare G.;
 RT "Phylogenetic inferences from chloroplast chlB gene sequences of
 RT Nephrolepis exaltata (Filicopsida), Ephedra altissima (Gnecopsida),
 RT and diverse land plants."
 RL Mol. Phylogenet. Evol. 6:19-29(1996).
 CC -1- FUNCTION: Uses Mg-ATP and reduced ferredoxin to reduce ring D of
 CC protochlorophyllide (Pchl) to form chlorophyllide a (Chl) (Chl) (Chl)
 CC (By similarity). This reaction is light-independent.
 CC -1- SUBUNIT: Light-independent chlorophyll biosynthesis.
 CC -1- SUBUNIT: Protochlorophyllide reductase is thought to be composed
 CC of three subunits: chlL, chlN and chlB. Could form a
 CC heterotrimer of two chlB and two chlN subunits.
 CC -1- SIMILARITY: Belongs to the chlB / chlN / chlB / chlN family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: U21315; AAC49429.1; -;
 DR EMBL: L25765; AAC37487.1; -;
 DR HAMAP: MF_00353; -; 1.
 DR InterPro: IPR000510; Oxred_nitrognase.
 DR InterPro: IPR005969; Protochl_reductB.
 DR Pfam: PF00148; oxidored_nitro; 1.
 DR TIGRFAMs: TIGR01278; DPOK_BchB; 1.
 DR Oxidoreductase; Photosynthesis; Chlorophyll biosynthesis; Chloroplast.
 KM SEQUENCE 509 AA; 57863 MW; E86032B894B4031 CRC64;

Query Match 65.3%; Score 32; DB 1; Length 509;
 Best Local Similarity 71.4%; Pred. No. 63;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 VDNRECI 9
 DB 249 IDNAECI 255

RESULT 29

LONI_SPIOL STANDARD; PRT; 875 AA.

AC 004979;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Lon protease homolog 1, mitochondrial precursor (EC 3.4.21.-).
 OS Spinacia oleracea (Spinach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Amaranthaceae; Spinacia.
 CC NCB1_TaxID=33562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=leaf;
 RA Inagaki N., Watanabe A., Satoh K.;
 RA Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix (Potential).
 CC -1- SIMILARITY: Belongs to peptidase family S16.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: D85610; BAA20482.1; -;
 DR PIR: T09142; T09142.
 DR MEROPS: S16.003; -;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003599; AAA_ATPase_cent.
 DR InterPro: IPR001270; Chaperin_c1pA/B.
 DR InterPro: IPR008269; Pept_S16_C.
 DR InterPro: IPR004815; Pept_S16_Ion.
 DR InterPro: IPR003111; Pept_S16_N.
 DR InterPro: IPR008268; Peptid_S16_AS.
 DR InterPro: IPR001984; Peptidase_S16.
 DR Pfam: PF00004; AAA; 1.
 DR Pfam: PF05362; Lon_C; 1.
 DR PRINTS: PR00300; CLPPROTEASEA.
 DR PRINTS: PR00830; ENDOLAPSE.
 DR SMART: SM00382; AAA; 1.
 DR SMART: SM00464; Lon; 1.
 DR TIGRFAMs: TIGR00763; lon; 1.
 DR PROSITE: PS01046; LON_SER; 1.
 DR Hydrolase; Serine protease; ATP-binding; Multigene family;
 KW Mitochondrion; Transit peptide.
 FT TRANSIT 1 875
 FT CHAIN ? 875
 FT NP BIND 409 416
 FT ACT SITE 784 784
 FT AC SITE 784 784
 FT BY SIMILARITY.
 SQ SEQUENCE 875 AA; 96782 MW; 7DD09003B0455A67 CRC64;

Query Match 65.3%; Score 32; DB 1; Length 875;
 Best Local Similarity 100.0%; Pred. No. 1; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVDNRE 7
 DB 652 GVDNRE 657

RESULT 30
 GPCR_LYMST STANDARD; PRT; 1115 AA.

AC P46023;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE G-protein coupled receptor GR1101 precursor.
 OS Lymnaea stagnalis (Great pond snail).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
 OC Lymnaeidae; Lymnaeidae; Lymnaea.
 CC NCB1_TaxID=6523;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=CNS;

RX MEDLINE=94255418; PubMed=8197140;
 RA Tensen C.P., van Kesteren E.R., Planta R.J., Cox K.J.A., Burke J.F.,
 RA van Heerikhuizen H., Vreugdenhil E.;
 RT "A G-protein-coupled receptor with low density lipoprotein-binding
 RT motifs suggests a role for lipoproteins in G-linked signal
 RT transduction."
 RL Proc. Natl. Acad. Sci. U.S.A. 91:4816-4820(1994).
 CC -1- FUNCTION: Might directly transduce signals carried by large
 CC extracellular (Lipo)protein(complex)es into neuronal events.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Predominantly expressed in a small number of
 CC neurons within the central nervous system and to a lesser extent
 CC in the heart.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC -1- SIMILARITY: Contains 12 LDL-receptor class A domains.
 CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z23104; CA80651.1; -
 DR PIR: S40241; S40241.
 DR HSSP: Q07954; 1CR8.
 DR InterPro: IPR000276; GPCR_Rhodopn.
 DR InterPro: IPR0002172; LDL_receptor_A.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR InterPro: IPR003591; LRR_typ.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00057; 1dl_recepte_a; 11.
 DR Pfam: PF00560; LRR; 6.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00237; GPCRHOOPSN.
 DR PRINTS: PR00261; LDLRECEPTOR.
 DR SMART: SM00192; LDla; 12.
 DR SMART: SM00369; LRR_TYP; 1.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
 DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
 DR PROSITE: PS01209; LDLa_1; 6.
 DR PROSITE: PS0068; LDLa_2; 11.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
 DR Leucine-rich repeat; Signal.
 KM SIGNAL 1 24
 FT CHAIN 25 1115
 FT DOMAIN 25 767
 FT TRANSMEM 768 788
 FT DOMAIN 789 801
 FT TRANSMEM 802 822
 FT DOMAIN 823 857
 FT TRANSMEM 858 878
 FT DOMAIN 879 887
 FT TRANSMEM 888 908
 FT DOMAIN 909 941
 FT TRANSMEM 942 962
 FT DOMAIN 963 988
 FT TRANSMEM 989 1009
 FT DOMAIN 1010 1017
 FT TRANSMEM 1018 1038
 FT DOMAIN 1039 1115
 FT TRANSMEM 1039 523
 FT DOMAIN 32 523
 FT TRANSMEM 32 523
 FT DOMAIN 36 79
 FT TRANSMEM 77 115
 FT DOMAIN 116 155
 FT TRANSMEM 156 196
 FT DOMAIN 195 232
 FT TRANSMEM 231 269
 FT TRANSMEM 231 269

FT DOMAIN 272 318
 FT DOMAIN 320 363
 FT DOMAIN 365 403
 FT DOMAIN 404 442
 FT DOMAIN 444 485
 FT DOMAIN 486 525
 FT REPEAT 557 579
 FT REPEAT 582 605
 FT REPEAT 607 629
 FT REPEAT 630 653
 FT REPEAT 655 677
 FT REPEAT 678 701
 FT REPEAT 703 725
 FT DISULFID 38 53
 FT DISULFID 46 66
 FT DISULFID 60 77
 FT DISULFID 79 91
 FT DISULFID 86 104
 FT DISULFID 98 113
 FT DISULFID 118 131
 FT DISULFID 138 153
 FT DISULFID 158 170
 FT DISULFID 165 183
 FT DISULFID 177 194
 FT DISULFID 202 220
 FT DISULFID 214 230
 FT DISULFID 233 245
 FT DISULFID 240 258
 FT DISULFID 252 267
 FT DISULFID 274 291
 FT DISULFID 282 294
 FT DISULFID 298 316
 FT DISULFID 332 339
 FT DISULFID 334 352
 FT DISULFID 346 361
 FT DISULFID 367 379
 FT DISULFID 374 392
 FT DISULFID 386 401
 FT DISULFID 406 418
 FT DISULFID 433 431
 FT DISULFID 425 440
 FT DISULFID 446 458
 FT DISULFID 453 474
 FT DISULFID 465 483
 FT DISULFID 488 500
 FT DISULFID 495 513
 FT DISULFID 507 523
 FT CARBOHYD 87 87
 FT CARBOHYD 166 166
 FT CARBOHYD 269 269
 FT CARBOHYD 318 318
 FT CARBOHYD 482 482
 FT CARBOHYD 502 502
 FT CARBOHYD 571 571
 FT CARBOHYD 618 618
 FT CARBOHYD 624 624
 FT CARBOHYD 685 685
 SQ SEQUENCE 1115 AA; 125865 MW; E9BB01297EEB356C CRC64;
 Query Match 65.3%; DB 1; Length 1115;
 Best Local Similarity 71.4%;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GYDNEC 8
 DB 107 GLDEREC 113
 RESULT 31
 M06_MOUSE
 ID MY06_MOUSE STANDARD; PRT; 1265 AA.
 AC Q64331;
 DT 16-OCT-2001 (rel. 40; Created)

approximately 50 years, affected individuals invariably have profound sensorineural deafness.

-1- DISEASE: Defects in MYO6 are the cause of autosomal recessive congenital neurosensory deafness type 37 (DFNB37) [MIM:607821].

-1- SIMILARITY: Contains 1 myosin-like globular head domain.

-1- SIMILARITY: Contains 1 IQ domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U09236; AAC51654.2; -

CC EMBL; AB002387; BAA20843.1; -

CC EMBL; AF229111; AAK00229.1; -

CC EMBL; AF229082; AAK00229.1; JOINED.

CC EMBL; AF229083; AAK00229.1; JOINED.

CC EMBL; AF229084; AAK00229.1; JOINED.

CC EMBL; AF229085; AAK00229.1; JOINED.

CC EMBL; AF229086; AAK00229.1; JOINED.

CC EMBL; AF229087; AAK00229.1; JOINED.

CC EMBL; AF229088; AAK00229.1; JOINED.

CC EMBL; AF229089; AAK00229.1; JOINED.

CC EMBL; AF229090; AAK00229.1; JOINED.

CC EMBL; AF229091; AAK00229.1; JOINED.

CC EMBL; AF229092; AAK00229.1; JOINED.

CC EMBL; AF229093; AAK00229.1; JOINED.

CC EMBL; AF229094; AAK00229.1; JOINED.

CC EMBL; AF229095; AAK00229.1; JOINED.

CC EMBL; AF229096; AAK00229.1; JOINED.

CC EMBL; AF229097; AAK00229.1; JOINED.

CC EMBL; AF229098; AAK00229.1; JOINED.

CC EMBL; AF229099; AAK00229.1; JOINED.

CC EMBL; AF229100; AAK00229.1; JOINED.

CC EMBL; AF229101; AAK00229.1; JOINED.

CC EMBL; AF229102; AAK00229.1; JOINED.

CC EMBL; AF229103; AAK00229.1; JOINED.

CC EMBL; AF229104; AAK00229.1; JOINED.

CC EMBL; AF229105; AAK00229.1; JOINED.

CC EMBL; AF229106; AAK00229.1; JOINED.

CC EMBL; AF229107; AAK00229.1; JOINED.

CC EMBL; AF229108; AAK00229.1; JOINED.

CC EMBL; AF229109; AAK00229.1; JOINED.

CC EMBL; AF229110; AAK00229.1; JOINED.

CC HSSP; P08799; 1MND.

CC Gene; HGNC:7605; MYO6.

CC MIM; 600970; -

CC MIM; 606346; -

CC MIM; 607821; -

CC GO; GO:0016461; C:unconventional myosin; TAS.

CC GO; GO:0003779; F:actin binding; TAS.

CC GO; GO:0005524; F:ATP binding; NAS.

CC GO; GO:0008307; F:structural constituent of muscle; NAS.

CC GO; GO:0006941; F:striated muscle contraction; NAS.

CC InterPro; IPR001609; IQ_region.

CC Pfam; PF00612; IQ_1.

CC PRINTS; PR00193; MYOSINHEAVY.

CC PRODOM; PD000355; myosin_head; 1.

CC SMART; SM00242; MYSC; 1.

CC PROSITE; PS50956; IQ; FALSE NEG.

CC Myosin; ATP-binding; Calmodulin-binding; Actin-binding; Coiled coil; Disease mutation; Deafness; Alternative splicing.

CC DOMAIN 1 759 MYOSIN HEAD-LIKE.

CC DOMAIN 834 IQ.

CC DOMAIN 848 1030 COILED COIL (POTENTIAL).

CC NP_BIND 151 158 ATP (POTENTIAL).

CC FT DOMAIN 665 672 ACTIN-BINDING (POTENTIAL).

CC FT VARSLIC 1037 1059 Missing (in isoform 2).

FT FT /FTId=VSP_007985.

FT VARIANT 216 216 E -> V (in DFNB37).

FT FT /FTId=VAR_016209.

FT VARIANT 442 442 C -> Y (in DFNB22).

FT FT /FTId=VAR_012110.

FT SEQUENCE 1285 AA; 148713 MW; BCB4FDE920712CD CRC64;

Query Match 65.3%; Score 32; DB 1; Length 1285;

Best Local Similarity 71.4%; Pred. No. 1.6e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNRECI 9

DB 509 VDNQDCI 515

RESULT 33

DYHC_EMBL

ID DYHC_EMBL STANDARD; PRT; 4344 AA.

AC P45444;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Dynein heavy chain, cytosolic (DYHC).

GN NUDA.

OS Emericella nidulans (Aspergillus nidulans).

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiales; Trichocommata; Emericella.

OX NCBI_TaxId=162425;

OX (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=94181539; PubMed=8134356;

RA Xiang X., Beckwith S.M., Morris R.N.;

RT "Cytosolic dynein is involved in nuclear migration in Aspergillus nidulans."

RL Proc. Natl. Acad. Sci. U.S.A. 91:2100-2104(1994).

CC -1- FUNCTION: Cytosolic dynein acts as a motor for the intracellular retrograde motility of vesicles and organelles along microtubules. Dynein has ATPase activity; the force-producing power stroke is thought to occur on release of ADP. Required to maintain uniform nuclear distribution in hyphae.

CC -1- SUBUNIT: Consists of at least two heavy chains and a number of intermediate and light chains.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem (which binds cargo and interacts with other dynein components), and the head or motor domain. The motor contains six tandemly-linked AAA domains in the head, which form a ring. A stalk-like structure (formed by two of the coiled coil domains) protrudes between AAA 4 and AAA 5 and terminates in a microtubule-binding site. A seventh domain may also contribute to this ring; it is not clear whether the N-terminus or the C-terminus forms this extra domain. There are four well-conserved and two non-conserved ATPase sites, one per AAA domain. Probably only one of these (within AAA 1) actually hydrolyzes ATP, the others may serve a regulatory function.

CC -1- SIMILARITY: Belongs to the dynein heavy chain family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U03904; AAA18338.1; -

CC PIR; A53489; A53489.

CC InterPro; IPR003593; AAA_ATPase.

CC InterPro; IPR004273; Dynein_heavy.

CC Pfam; PF03028; Dynein_heavy; 1.

CC SMART; SM00382; AAA; 3.

CC Motor protein; Dynein; Microtubule; ATP-binding; Repeat; Coiled coil.

```

FT DOMAIN 1 1894 STEM (BY SIMILARITY).
FT DOMAIN 1695 2120 AAA 1 (BY SIMILARITY).
FT DOMAIN 2190 2448 AAA 2 (BY SIMILARITY).
FT DOMAIN 2553 2800 AAA 3 (BY SIMILARITY).
FT DOMAIN 2894 3163 AAA 4 (BY SIMILARITY).
FT DOMAIN 3179 3463 STALK (BY SIMILARITY).
FT DOMAIN 3549 3780 AAA 5 (BY SIMILARITY).
FT DOMAIN 3994 4206 AAA 6 (BY SIMILARITY).
FT DOMAIN 667 687 COILED COIL (POTENTIAL).
FT DOMAIN 913 933 COILED COIL (POTENTIAL).
FT DOMAIN 1321 1341 COILED COIL (POTENTIAL).
FT DOMAIN 1548 1564 COILED COIL (POTENTIAL).
FT DOMAIN 1627 1656 COILED COIL (POTENTIAL).
FT DOMAIN 3179 3270 COILED COIL (POTENTIAL).
FT DOMAIN 3407 3463 COILED COIL (POTENTIAL).
FT DOMAIN 3700 3798 COILED COIL (POTENTIAL).
FT NP_BIND 1933 1940 ATP (POTENTIAL).
FT NP_BIND 2223 2230 ATP (POTENTIAL).
FT NP_BIND 2592 2599 ATP (POTENTIAL).
FT NP_BIND 2932 2939 ATP (POTENTIAL).
SO SEQUENCE 4344 AA; 492470 MW; 302A586C4F8BD019 CRC64;

Query Match 65.3%; Score 32; DB 1; Length 4344;
Best Local Similarity 62.5%; Pred. No. 5.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GVDNRECT 9
Db 4271 GVDNDECL 4278

RESULT 34
YSH2 CAEEL STANDARD; PRT; 160 AA.
ID YSH2 CAEEL
AC Q093944;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 18.4 kDa protein Fl2A10.2 in chromosome II.
GN Fl2A10.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Geisler C.;
RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: STRONG, IN THE C-TERMINUS, WITH C.ELEGANS ZK177.2 AND
CC Fl2A10.6.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. By non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U28731; AAA68295.1; -
DR PIR; T16043; T16043.
DR WormPep; Fl2A10.2; CE01905.
KW Hypothetical protein.
SQ SEQUENCE 160 AA; 18439 MW; C3763C8C8C76B2EC CRC64;

Query Match 63.3%; Score 31; DB 1; Length 160;
Best Local Similarity 62.5%; Pred. No. 31;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GVDNRECT 9
Db 18 GVDNRECT 25

```

```

RESULT 35
PYRE CAUCR STANDARD; PRT; 194 AA.
ID PYRE CAUCR
AC Q9A810;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Orotate phosphoribosyltransferase (EC 2.4.2.10) (OPRT) (OPRase).
GN PYRE OR CC1555.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
CC NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Niernan W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Ueberback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate + diphosphate = orotate
CC + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -1- COFACTOR: Magnesium (by similarity).
CC -1- PATHWAY: Pyrimidine biosynthesis; fifth step.
CC -1- SUBUNIT: Homodimer (by similarity).
CC -1- SIMILARITY: Belongs to the purine/pyrimidine
CC phosphoribosyltransferase family. PYRE subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. By non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE005830; AAK23534.1; -
DR PIR; B87442; B87442.
DR TIGR; CC1555; -
DR HAMAP; MF_01208; -; 1.
DR InterPro; IPR002375; Pr/DY_ip.transf.
DR InterPro; IPR000836; PRTtransferase.
DR InterPro; IPR006273; PYRE_chem.
DR Pfam; PF00156; Pribosyltran; 1.
DR TIGRFAMs; TIGR01367; PYRE_Therm; 1.
DR PROSITE; PS00103; PUR_PYR_PR_TRANSFR; 1.
KW Pyrimidine biosynthesis; Transferase; Glycosyltransferase; Magnesium;
KW Complete proteome.
FT DOMAIN 112 124 PRPP-BINDING (BY SIMILARITY).
SQ SEQUENCE 194 AA; 20583 MW; FB58C09C973FBA83 CRC64;

Query Match 63.3%; Score 31; DB 1; Length 194;
Best Local Similarity 62.5%; Pred. No. 38;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GVDNRECT 9
Db 122 GLSSRECT 129

RESULT 36
HIS1 STAM STANDARD; PRT; 204 AA.
ID HIS1 STAM
AC Q99QW2;
DT 28-FEB-2003 (Rel. 41, Created)

```

DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE ATP phosphoribosyltransferase (EC 2.4.2.17).
 GN HISG OR SAV2679 OR SA2471 OR MM2598.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
 OS Staphylococcus aureus (strain N315), and
 OS Staphylococcus aureus (strain MM2).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 NCBI_TaxID=158879, 158879, 196620;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mu50 / ATCC 700699, and N315;
 RX MEDLINE=21311952; PubMed=1181846;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian U.-Q., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hackett M., Ogasawara N., Hayashi H., Hiramatsu K.,
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 RT aureus";
 RL Lancet 357:1225-1240(2001).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MM2;
 RX MEDLINE=22040717; PubMed=12044378;
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
 RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
 RA Yamamoto K., Hiramatsu K.;
 RT "Genome and virulence determinants of high virulence community-
 RT acquired MRSA";
 RL Lancet 359:1819-1827(2002).
 CC -1- CATALYTIC ACTIVITY: 1-(5-phospho-D-ribose)-ATP + diphosphate =
 CC ATP + 5-phospho-alpha-D-ribose 1-diphosphate.
 CC -1- PATHWAY: Histidine biosynthesis; first step. Very important in the
 CC regulation of histidine metabolism.
 CC -1- SUBUNIT: Homohexamer (By similarity). Binds to hisZ possibly to
 CC allow the regulation of hisG transferase activity by histidine (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- DOMAIN: Lacks the C-terminal regulatory region which is replaced
 CC by hisZ.
 CC -1- SIMILARITY: Belongs to the ATP phosphoribosyltransferase family.
 CC Short subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AP003366; BAB58841.1; -;
 CC EMBL: AP003366; BAB58841.1; -;
 CC EMBL: AP004631; BAB96463.1; -;
 CC PIR: G90076; G90076.
 CC HAMAP: MF_01018; -; 1.
 CC InterPro: IPR001348; ATP_phospho_trans.
 CC Pfam: PF01634; HisG; 1.
 CC ProDom: PD003516; ATP_phospho_trans; 1.
 CC PROSITE: PS01316; ATP_P_PHOSRIBOSYLTR; FALSE NEG.
 CC Histidine biosynthesis; transferase; Glycosyltransferase;
 CC Complete proteome.
 CC KW Complete proteome.
 CC SQ SEQUENCE 204 AA; 22643 MW; EB81C7225FDB18C2 CRC64;
 CC
 CC Query Match 63.3%; Score 31; DB 1; Length 204;
 CC Best Local Similarity 85.7%; Pred. No. 40;
 CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC 3 VDNRECT 9
 CC |||||

Db 42 VDNRECT 48
 RESULT 37
 ID RFAP_ECOLI STANDARD; PRT; 265 AA.
 AC P25741;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Lipopolysaccharide core biosynthesis protein rfaP.
 GN RFAP OR WAP OR B3630.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 NCBI_TaxID=562;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KL2;
 RX MEDLINE=92121133; PubMed=1732225;
 RA Parker C.T., Pradel E., Schnaitman C.A.;
 RT "Identification and sequences of the lipopolysaccharide core
 RT biosynthetic genes rfaQ, rfaP, and rfaG of Escherichia coli K-12";
 RL J. Bacteriol. 174:930-934(1992).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KL2 / MG1655;
 RX MEDLINE=94316500; PubMed=8041620;
 RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes";
 RL Nucleic Acids Res. 22:2576-2586(1994).
 CC -1- FUNCTION: INVOLVED IN ATTACHMENT OF PHOSPHATE-CONTAINING
 CC SUBSTITUENTS TO THE INNER CORE.
 CC -1- PATHWAY: Lipopolysaccharide core biosynthesis.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M80599; AAA24083.1; -;
 CC EMBL: U00039; AAB18607.1; -;
 CC EMBL: AE000440; AAC7654.1; -;
 CC PIR: C42595; C42595.
 CC EcoGene: EG11340; rfaP.
 CC InterPro: IPR007924; WAP.
 CC Pfam: PF05260; WAP; 1.
 CC Lipopolysaccharide biosynthesis; Transferase; Complete proteome.
 CC KW Lipopolysaccharide biosynthesis; Transferase; Complete proteome.
 CC SQ SEQUENCE 265 AA; 30872 MW; E77F1B935ABDBD CRC64;
 CC
 CC Query Match 63.3%; Score 31; DB 1; Length 265;
 CC Best Local Similarity 50.0%; Pred. No. 52;
 CC Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC 1 AGVDNREC 8
 CC |||||
 CC Db 156 AGINHRDC 163
 CC
 CC RESULT 38
 CC CMCH STRCL STANDARD; PRT; 521 AA.
 AC O85728;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DE 3'-hydroxymethylcephem-O-carbamoyltransferase (EC 2.1.3.-)
 DE (3'-hydroxymethylcephem-O-CASE) (CCT).
 GN CMCH.

```

OS Streptomyces clavuligerus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OK NCBI_TaxID=1901;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27064 / DSM 738 / NRRL 3585;
RX MEDLINE=98361893; PubMed=9696752;
RA Alexander D.C., Jensen S.E.;
RT "Investigation of the Streptomyces clavuligerus cephamycin C gene
cluster and its regulation by the Ccar protein.";
RL J. Bacteriol. 180:4068-4079(1998).
CC -1- FUNCTION: CATALYZES THE CARBAMOYLATION REACTION IN THE CEPHAMYCIN
CC BIOSYNTHESIS.
CC -1- PATHWAY: Cephamycin biosynthesis.
CC -1- SIMILARITY: Belongs to the nodU/cmh family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF073897; AAC2493.1; -
DR InterPro; IPR003696; Carbtransf.
DR Pfam; PF02543; Cmh_NodU; 1.
KW Antibiotic biosynthesis; Transferase.
SQ SEQUENCE 521 AA; 57369 MW; A80C078999B137B1 CRC64;

Query Match 63.3%; Score 31; DB 1; Length 521;
Best Local Similarity 62.5%; Pred. No. 1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGVDNREC 8
DB 229 AGVESEC 236

RESULT 39
CP11_ONCMY STANDARD; PRT; 522 AA.
AC Q92110; O42195; P10609; P79830;
DT 01-JUL-1989 (Rel. 11, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 1A1 (EC 1.14.14.1) (CYP1A1) (CYP1A2).
GN CYP1A1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Procyanocephalygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94213487; PubMed=8161204;
RA Berndtson A.K., Chen T.T.;
RT "Two unique CYP1 genes are expressed in response to 3-
RT methylcholanthrene treatment in rainbow trout.";
RL Arch. Biochem. Biophys. 310:187-195(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Bailey G., You L., Harttig U.;
RT "Cloning, sequencing and functional expression of two trout CYP1A
RT cDNAs in yeast.";
RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Bailey G., You L., Harttig U.;
RT "Cloning, sequencing and alicoxin B1 metabolism by multiple rainbow
RT trout CYP1A cDNAs expressed in yeast.";

```

```

RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89078171; PubMed=3203599;
RA Hellmann L.U., Sheen Y.-Y., Bigelow S.W., Nebert D.W.;
RT "Trout P4501A1: cDNA and deduced protein sequence, expression in
RT liver, and evolutionary significance.";
RL DNA 7:379-387(1988).
CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. They oxidize a variety of structurally unrelated
CC compounds, including steroids, fatty acids, and xenobiotics.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- TISSUE SPECIFICITY: Liver.
CC -1- INDUCTION: By 3-methylcholanthrene (3MC).
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC -1- CAUTION: THE SEQUENCE FROM REF.4 WAS A CHIMERA: ITS N-TERMINAL
CC PART HAS BEEN SHOWN TO BE DERIVED FROM WHAT IS NOW KNOWN AS THE
CC CYP1A3. CYP1A1 HAS ALSO BEEN CALLED CYP1A2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S69278; AAD14036.1; -
DR EMBL; U62797; AAB40627.1; -
DR EMBL; AF015660; AAB69383.1; -
DR EMBL; M21310; AAA49550.1; ALT_SEQ.
DR PIR; A28789; A28789.
DR PIR; S5157; S5157.
DR HSSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008066; BP4501a.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01683; EP4501CYP1A.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 463 463 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT METAL 51 51 L -> P (IN REF. 3).
FT CONFLICT 97 97 V -> D (IN REF. 3).
FT CONFLICT 124 124 K -> N (IN REF. 3).
FT CONFLICT 306 306 Q -> H (IN REF. 3).
FT CONFLICT 390 390 T -> I (IN REF. 3).
FT CONFLICT 459 459 D -> G (IN REF. 2).
FT CONFLICT 470 470 R -> H (IN REF. 3).
FT CONFLICT 474 474 Y -> F (IN REF. 1).
FT CONFLICT 483 483 R -> K (IN REF. 2).
SQ SEQUENCE 522 AA; 59344 MW; 9D063B5891102CE9 CRC64;

Query Match 63.3%; Score 31; DB 1; Length 522;
Best Local Similarity 62.5%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVDNRECI 9
DB 457 GMDRRCTI 464

RESULT 40
URE2_HELPF STANDARD; PRT; 569 AA.
AC Q08716;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE  Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN  URER.
OS  Helicobacter felis.
OC  Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC  Helicobacteriaceae; Helicobacter.
OX  NCBI_TaxId=214;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=94018627; PubMed=8412683;
RA  Ferrero R.L., Labigne A.;
RT  "Cloning, expression and sequencing of Helicobacter felis urease
RT  genes.";
RL  Mol. Microbiol. 9:323-333(1993).
CC  -1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC  -1- COPACTOR: Binds 2 nickel ions per subunit (by similarity).
CC  -1- SUBUNIT: CONSIST OF TWO SUBUNITS (ALPHA AND BETA).
CC  -1- PTM: Lys-219 is carbamylated. The carbamoyl group provides the
CC  ligands for the two nickel ions (by similarity).
CC  -1- SIMILARITY: Belongs to the urease family.
CC  -1- CAUTION: IN HELICOBACTER THE BETA SUBUNIT IS WHAT IS KNOWN, IN
CC  OTHER BACTERIA AS THE ALPHA SUBUNIT.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X69080; CAA48826.1; -
DR  PIR; S35291; S35291.
DR  HSP; F18314; IFWE.
DR  InterPro; IPR006680; Amidohydro_1.
DR  InterPro; IPR005848; Pept_M38_urease.
DR  InterPro; IPR008295; Urease_alpha.
DR  Pfam; PF01979; Amidohydro_1; 1.
DR  Pfam; PRO0449; urease; 1.
DR  Pfam; PRO0449; urease; 1.
DR  PRINTS; PRSF001226; Urease_alpha; 1.
DR  PRINTS; PRO1752; UREASE.
DR  PROSITE; PS00145; UREASE_2; 1.
DR  PROSITE; PS01120; UREASE_1; 1.
KW  Hydrolase; Metal-binding; Nickel.
FT  METAL 136 136 NICKEL 2 (BY SIMILARITY).
FT  METAL 138 138 NICKEL 2 (BY SIMILARITY).
FT  METAL 219 219 NICKEL 1 AND 2 (BY SIMILARITY).
FT  METAL 248 248 NICKEL 1 (BY SIMILARITY).
FT  METAL 274 274 NICKEL 1 (BY SIMILARITY).
FT  METAL 362 362 NICKEL 2 (BY SIMILARITY).
FT  METAL 362 362 NICKEL 2 (BY SIMILARITY).
FT  ACT_SITE 322 322 BY SIMILARITY.
SQ  SEQUENCE 569 AA; 61702 MW; 66DBB99E2CC7CEA3 CRC64;

Query Match 63.3%; Score 31; DB 1; Length 569;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVDNRECT 9
Db 107 GVDNNLCV 114

```

Search completed: March 3, 2004, 10:23:20
 Job time : 11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 09:48:23 ; Search time 36.5 Seconds
(without alignments)
77.799 Million cell updates/sec

Title: US-09-980-064-2
Perfect score: 49
Sequence: 1 AGVDNRECI 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	419	12	Q81004 human papill
2	49	100.0	494	12	Q81010 human papill
3	49	100.0	494	12	Q81002 human papill
4	49	100.0	494	12	Q81003 human papill
5	49	100.0	494	12	Q81006 human papill
6	49	100.0	494	12	Q81011 human papill
7	49	100.0	494	12	Q81005 human papill
8	49	100.0	494	12	Q80998 human papill
9	49	100.0	494	12	Q81004 human papill
10	49	100.0	494	12	Q81007 human papill
11	49	100.0	494	12	Q80996 human papill
12	49	100.0	494	12	Q80997 human papill
13	49	100.0	494	12	Q81008 human papill
14	49	100.0	494	12	Q80999 human papill
15	49	100.0	494	12	Q81009 human papill
16	49	100.0	494	12	Q81012 human papill

17	49	100.0	494	12	Q81001 human papill
18	49	100.0	494	12	Q90204 human papill
19	49	100.0	494	12	Q81000 human papill
20	49	100.0	495	12	Q12343 human papill
21	49	100.0	505	12	Q9WH57 human papill
22	49	100.0	505	12	Q9WLO5 human papill
23	49	100.0	505	12	Q9WH56 human papill
24	49	100.0	505	12	Q81015 human papill
25	49	100.0	505	12	Q9WPH4 human papill
26	49	100.0	505	12	Q92282 human papill
27	49	100.0	505	12	Q9WH55 human papill
28	49	100.0	505	12	Q9WLO6 human papill
29	49	100.0	531	12	Q00530 human papill
30	49	100.0	531	12	Q8B5N9 human papill
31	49	100.0	531	12	Q8B5N5 human papill
32	49	100.0	531	12	Q80ML8 human papill
33	41	83.7	504	12	Q81017 human papill
34	39	79.6	529	12	Q90730 human papill
35	36	73.5	310	16	Q7UKH0 rhodopsin
36	36	73.5	445	16	Q9AMT8 bradyrhizob
37	36	73.5	474	5	Q93879 caenorhabdi
38	36	73.5	539	5	Q95R86 drosophila
39	36	73.5	1039	10	Q9M924 arabidopsis
40	35	71.4	243	16	Q8Z640 salmone
41	35	71.4	444	2	Q93H51 bradyrhizob
42	35	71.4	501	2	Q83XA3 streptomyces
43	35	71.4	610	2	Q9ACT6 streptomyces
44	35	71.4	840	2	Q83W6 streptomyces
45	35	71.4	879	16	Q9AD14 streptomyces

ALIGNMENTS

RESULT 1
ID Q81004 PRELIMINARY: PRT: 419 AA.

AC Q81004: 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE L1 protein (Fragment).
LN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang J., Chen R.;
RT "Cloning and sequence analysis of Human papillomavirus type 16 L1 gene in China."
RT Submitted (May-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF512011; AM74159.1; -
DR GO: GO:0019028; C: viral capsid; IEA.
DR GO: GO:0005198; F: structural molecule activity; IEA.
DR InterPro: IPR002210; PV_capsid_L1.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00500; late_protein_L1; 1.
DR PRINTS: PR00865; HPVcapsid_L1.
DR ProDom: PD000544; PV_capsid_L1; 1.
FT NON TER 419
SQ SEQUENCE 419 AA; 46880 MW; 82099A0D5F584560 CRC64;

Query Match 100.0%; Score 49; DB 12; Length 419;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGVDNRECI 9
Db 164 AGVDNRECI 172

```

RESULT 2
Q81010 PRELIMINARY; PRT; 494 AA.
ID Q81010
AC Q81010;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 7632;
RX MEDLINE=96079021; Pubmed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
  Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
  populations characterized by nucleotide sequence analysis of the E6,
  L2, and L1 coding segments.";
  J. Virol. 69:7743-7753(1995).
  [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 7632;
RA Farmer A.D.;
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U34189; AAA91736.1; -.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 494
SQ SEQUENCE 494 AA; 55094 MW; ECB31B09B8EFBAE3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 49; DB 12; Length 494;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
Db 130 AGVDNRECI 138

RESULT 3
Q81002 PRELIMINARY; PRT; 494 AA.
ID Q81002
AC Q81002;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 3136;
RX MEDLINE=96079021; Pubmed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
  Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
  populations characterized by nucleotide sequence analysis of the E6,
  L2, and L1 coding segments.";
  J. Virol. 69:7743-7753(1995).
  [2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=OR 3136;
RA Farmer A.D.;
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U34174; AAA91721.1; -.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 494
SQ SEQUENCE 494 AA; 55090 MW; 36335F09BCAFF3CE CRC64;

Query Match
Best Local Similarity 100.0%; Score 49; DB 12; Length 494;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
Db 130 AGVDNRECI 138

RESULT 4
Q81003 PRELIMINARY; PRT; 494 AA.
ID Q81003
AC Q81003;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 3759;
RX MEDLINE=96079021; Pubmed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
  Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
  populations characterized by nucleotide sequence analysis of the E6,
  L2, and L1 coding segments.";
  J. Virol. 69:7743-7753(1995).
  [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 3759;
RA Farmer A.D.;
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U34176; AAA91723.1; -.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 494
SQ SEQUENCE 494 AA; 55100 MW; 0BF1732B2B831233 CRC64;

Query Match
Best Local Similarity 100.0%; Score 49; DB 12; Length 494;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
Db 130 AGVDNRECI 138

```



```
RESULT 5
ID Q81006 PRELIMINARY; PRT; 494 AA.
AC Q81006;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
LN 1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 6106;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
  Jensen S.A.;
  "Human papillomavirus type 16 variant lineages in United States
  populations characterized by nucleotide sequence analysis of the E6,
  L2, and L1 coding segments.";
  J. Virol. 69:7743-7753(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 6106;
RA Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
DR GO: GO:0019028; C:viral capsid, IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002210; PV_capsid_L1.
DR Pfam: PF00500; late_protein_L1; 1.
DR PRINTS: PR00865; HPVcapsidL1.
DR ProDom: PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 494
FT SEQUENCE 494 AA; 55080 MW; EBC46B09EBEAB1E3 CRC64;
SQ

Query Match 100.0%; Score 49; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGVDNRECT 9
Db 130 AGVDNRECT 138

RESULT 6
ID Q81011 PRELIMINARY; PRT; 494 AA.
AC Q81011;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
LN 1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 8160;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
  Jensen S.A.;
  "Human papillomavirus type 16 variant lineages in United States
  populations characterized by nucleotide sequence analysis of the E6,
  L2, and L1 coding segments.";
  J. Virol. 69:7743-7753(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 8160;
RA Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
DR GO: GO:0019028; C:viral capsid, IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002210; PV_capsid_L1.
DR Pfam: PF00500; late_protein_L1; 1.
DR PRINTS: PR00865; HPVcapsidL1.
DR ProDom: PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 494
FT SEQUENCE 494 AA; 55051 MW; BD92AC39CB68D2CE CRC64;
SQ

Query Match 100.0%; Score 49; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGVDNRECT 9
Db 130 AGVDNRECT 138
```

```
RESULT 7
ID Q81005 PRELIMINARY; PRT; 494 AA.
AC Q81005;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
LN 1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 5110;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
  Jensen S.A.;
  "Human papillomavirus type 16 variant lineages in United States
  populations characterized by nucleotide sequence analysis of the E6,
  L2, and L1 coding segments.";
  J. Virol. 69:7743-7753(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 5110;
RA Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
DR GO: GO:0019028; C:viral capsid, IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002210; PV_capsid_L1.
DR Pfam: PF00500; late_protein_L1; 1.
DR PRINTS: PR00865; HPVcapsidL1.
DR ProDom: PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 494
FT SEQUENCE 494 AA; 55063 MW; 735F7C2B86CAC58F CRC64;
SQ

Query Match 100.0%; Score 49; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGVDNRECT 9
Db 130 AGVDNRECT 138

RESULT 8
ID Q81005 PRELIMINARY; PRT; 494 AA.
AC Q81005;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
LN 1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 5110;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
  Jensen S.A.;
  "Human papillomavirus type 16 variant lineages in United States
  populations characterized by nucleotide sequence analysis of the E6,
  L2, and L1 coding segments.";
  J. Virol. 69:7743-7753(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 5110;
RA Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
DR GO: GO:0019028; C:viral capsid, IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002210; PV_capsid_L1.
DR Pfam: PF00500; late_protein_L1; 1.
DR PRINTS: PR00865; HPVcapsidL1.
DR ProDom: PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 494
FT SEQUENCE 494 AA; 55051 MW; BD92AC39CB68D2CE CRC64;
SQ

Query Match 100.0%; Score 49; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGVDNRECT 9
Db 130 AGVDNRECT 138
```

```
RESULT 8
Q80998      PRELIMINARY;      PRT;      494 AA.
AC  Q80998;
DT  01-NOV-1996 (TREMBlrel. 01, Created)
DT  01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT  01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE  Late major capsid protein L1 (Fragment).
GN  L1.
OS  Human papillomavirus.
OC  Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC  Papillomavirus.
OX  NCBI_TaxID=10566;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=NM 1197;
RX  MEDLINE=96079021; PubMed=7494284;
RA  Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RT  "Human papillomavirus type 16 variant lineages in United States
RT  populations characterized by nucleotide sequence analysis of the E6,
RT  L2, and L1 coding segments.";
RL  J. Virol. 69:7743-7753 (1995).
[2]
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=NM 1197;
RA  Farmer A.D.;
RA  Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR  EMBL; U34167; AAA91714.1; -.
DR  GO; GO:0019028; C:Viral capsid; IEA.
DR  GO; GO:0005198; F:Structural molecule activity; IEA.
DR  InterPro; IPR002210; PV_capsid_L1.
DR  InterPro; IPR008975; Viral_cap_coat.
DR  Pfam; PF00500; late_protein_L1; 1.
DR  PRINTS; PR00865; HPVcapsidL1.
DR  Prodom; PD000544; PV_capsid_L1; 1.
FT  NON_TER 1 494
SQ  SEQUENCE 494 AA; 55021 MW; E446FD38994FB3DD CRC64;

Query Match      100.0%; Score 49; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 AGVDNRECI 9
Db  130 AGVDNRECI 138

RESULT 9
Q81004      PRELIMINARY;      PRT;      494 AA.
AC  Q81004;
DT  01-NOV-1996 (TREMBlrel. 01, Created)
DT  01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT  01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE  Late major capsid protein L1 (Fragment).
GN  L1.
OS  Human papillomavirus.
OC  Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC  Papillomavirus.
OX  NCBI_TaxID=10566;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=OR 4541;
RX  MEDLINE=96079021; PubMed=7494284;
RA  Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RT  "Human papillomavirus type 16 variant lineages in United States
RT  populations characterized by nucleotide sequence analysis of the E6,
RT  L2, and L1 coding segments.";
```

```
RL  J. Virol. 69:7743-7753 (1995).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=OR 4541;
RA  Farmer A.D.;
RA  Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR  EMBL; U34177; AAA91724.1; -.
DR  GO; GO:0019028; C:Viral capsid; IEA.
DR  GO; GO:0005198; F:Structural molecule activity; IEA.
DR  InterPro; IPR002210; PV_capsid_L1.
DR  InterPro; IPR008975; Viral_cap_coat.
DR  Pfam; PF00500; late_protein_L1; 1.
DR  PRINTS; PR00865; HPVcapsidL1.
DR  Prodom; PD000544; PV_capsid_L1; 1.
FT  NON_TER 1 494
SQ  SEQUENCE 494 AA; 55063 MW; 735F7C2B86CACS8F CRC64;

Query Match      100.0%; Score 49; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 AGVDNRECI 9
Db  130 AGVDNRECI 138

RESULT 10
Q81007      PRELIMINARY;      PRT;      494 AA.
AC  Q81007;
DT  01-NOV-1996 (TREMBlrel. 01, Created)
DT  01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT  01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE  Late major capsid protein L1 (Fragment).
GN  L1.
OS  Human papillomavirus.
OC  Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC  Papillomavirus.
OX  NCBI_TaxID=10566;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=OR 6170;
RX  MEDLINE=96079021; PubMed=7494284;
RA  Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RT  "Human papillomavirus type 16 variant lineages in United States
RT  populations characterized by nucleotide sequence analysis of the E6,
RT  L2, and L1 coding segments.";
RL  J. Virol. 69:7743-7753 (1995).
[2]
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=OR 6170;
RA  Farmer A.D.;
RA  Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR  EMBL; U34184; AAA91731.1; -.
DR  GO; GO:0019028; C:Viral capsid; IEA.
DR  GO; GO:0005198; F:Structural molecule activity; IEA.
DR  InterPro; IPR002210; PV_capsid_L1.
DR  InterPro; IPR008975; Viral_cap_coat.
DR  Pfam; PF00500; late_protein_L1; 1.
DR  PRINTS; PR00865; HPVcapsidL1.
DR  Prodom; PD000544; PV_capsid_L1; 1.
FT  NON_TER 1 494
SQ  SEQUENCE 494 AA; 55051 MW; BD92AC39CB6BD2CE CRC64;

Query Match      100.0%; Score 49; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 AGVDNRECI 9
Db  130 AGVDNRECI 138
```

DB 130 AGVDNRECI 138

RESULT 11

080996 PRELIMINARY; PRT; 494 AA.

ID 080996; AC 080996;

DR 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

LN Late major capsid protein l1 (Fragment).

GN l1.

OS Human papillomavirus.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.

OX NCBI_TaxID=10566;

RP SEQUENCE FROM N.A.

RC STRAIN=NM_4094;

RX MEDLINE=96079021; PubMed=7494284;

RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A., Jenison S.A.;

RT "Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, l2, and l1 coding segments.";

RL J. Virol. 69:7743-7753(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=NM_4094;

RA Farmer A.D.;

RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.

DR EMBL; U34165; AA91712.1; -.

DR GO; GO:0019028; C:viral capsid; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR002210; PV_capsid_l1.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00500; late_protein_l1; 1.

DR PRINTS; PR00865; HPVcapsidl1.

DR ProDom; PD000544; PV_capsid_l1; 1.

FT NON_TER 1

FT NON_TER 494

SQ SEQUENCE 494 AA; 55063 MW; 735F7C2B86CAC58F CRC64;

Query Match 100.0%; Score 49; DB 12; Length 494;

Best Local Similarity 100.0%; Pred. No. 0.097;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9

DB 130 AGVDNRECI 138

RESULT 12

080997 PRELIMINARY; PRT; 494 AA.

ID 080997; AC 080997;

DR 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

LN Late major capsid protein l1 (Fragment).

GN l1.

OS Human papillomavirus.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.

OX NCBI_TaxID=10566;

RP SEQUENCE FROM N.A.

RC STRAIN=NM_9999;

RX MEDLINE=96079021; PubMed=7494284;

RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A., Jenison S.A.;

RT "Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6,

RT l2, and l1 coding segments.";

RL J. Virol. 69:7743-7753(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=NM_9999;

RA Farmer A.D.;

RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.

DR EMBL; U34165; AA91713.1; -.

DR GO; GO:0019028; C:viral capsid; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR002210; PV_capsid_l1.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00500; late_protein_l1; 1.

DR PRINTS; PR00865; HPVcapsidl1.

DR ProDom; PD000544; PV_capsid_l1; 1.

FT NON_TER 1

FT NON_TER 494

SQ SEQUENCE 494 AA; 55063 MW; 735F7C2B86CAC58F CRC64;

Query Match 100.0%; Score 49; DB 12; Length 494;

Best Local Similarity 100.0%; Pred. No. 0.097;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9

DB 130 AGVDNRECI 138

RESULT 13

081008 PRELIMINARY; PRT; 494 AA.

ID 081008; AC 081008;

DR 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

LN Late major capsid protein l1 (Fragment).

GN l1.

OS Human papillomavirus.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.

OX NCBI_TaxID=10566;

RP SEQUENCE FROM N.A.

RC STRAIN=OR_7574;

RX MEDLINE=96079021; PubMed=7494284;

RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A., Jenison S.A.;

RT "Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, l2, and l1 coding segments.";

RL J. Virol. 69:7743-7753(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=OR_7574;

RA Farmer A.D.;

RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.

DR EMBL; U34187; AA91734.1; -.

DR GO; GO:0019028; C:viral capsid; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR002210; PV_capsid_l1.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00500; late_protein_l1; 1.

DR PRINTS; PR00865; HPVcapsidl1.

DR ProDom; PD000544; PV_capsid_l1; 1.

FT NON_TER 1

FT NON_TER 494

SQ SEQUENCE 494 AA; 55008 MW; E55B8D4EF109F59E CRC64;

Query Match 100.0%; Score 49; DB 12; Length 494;

Best Local Similarity 100.0%; Pred. No. 0.097;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9

Db 130 AGVDNRECI 138

RESULT 14

```

O80999          PRELIMINARY;      PRT;      494 AA.
ID  O80999;
AC  O80999;
DT  01-NOV-1996 (TREMblrel. 01, Created)
DT  01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT  01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE  Late major capsid protein L1 (Fragment).
GN  L1.
OS  Human papillomavirus.
OC  Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC  Papillomavirus.
OX  NCBI_TaxID=10566;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=NM_T455;
RX  MEDLINE=96079021; PubMed=7494284;
RA  Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA  Jenison S.A.;
RT  "Human papillomavirus type 16 variant lineages in United States
RT  populations characterized by nucleotide sequence analysis of the E6,
RT  L2, and L1 coding segments.";
RT  J. Virol. 69:7743-7753 (1995).
RN  (2)
RN  SEQUENCE FROM N.A.
RP  STRAIN=NM_T455;
RC  Farmer A.D.;
RA  Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR  EMBL; U34163; AAA91716.1; -.
DR  GO; GO:0019028; C:Viral capsid; IEA.
DR  GO; GO:0005198; F:Structural molecule activity; IEA.
DR  InterPro; IPR002210; PV_capsid_L1.
DR  InterPro; IPR008975; Viral_cap_coat.
DR  Pfam; PF00500; late_protein_L1; 1.
DR  PRINTS; PR00865; HPVcapsidL1.
DR  ProDom; PD000544; PV_capsid_L1; 1.
FT  NON_TER 1
FT  NON_TER 494
SQ  SEQUENCE 494 AA; 55051 MW; 8DE5FC4CCD7DE5D8 CRC64;

Query Match          100.0%; Score 49; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 AGVDNRECI 9
    |||||
    |||||
    |||||
Db  130 AGVDNRECI 138

RESULT 15
O81009          PRELIMINARY;      PRT;      494 AA.
ID  O81009;
AC  O81009;
DT  01-NOV-1996 (TREMblrel. 01, Created)
DT  01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT  01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE  Late major capsid protein L1 (Fragment).
GN  L1.
OS  Human papillomavirus.
OC  Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC  Papillomavirus.
OX  NCBI_TaxID=10566;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=OR_7587;
RX  MEDLINE=96079021; PubMed=7494284;
RA  Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA  Jenison S.A.;
RT  "Human papillomavirus type 16 variant lineages in United States
RT  populations characterized by nucleotide sequence analysis of the E6,
RT  L2, and L1 coding segments.";
RT  J. Virol. 69:7743-7753 (1995).
RN  (2)
RN  SEQUENCE FROM N.A.
RP  STRAIN=OR_7587;
RC  Farmer A.D.;
RA  Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR  EMBL; U34163; AAA91716.1; -.
DR  GO; GO:0019028; C:Viral capsid; IEA.
DR  GO; GO:0005198; F:Structural molecule activity; IEA.
DR  InterPro; IPR002210; PV_capsid_L1.
DR  InterPro; IPR008975; Viral_cap_coat.
DR  Pfam; PF00500; late_protein_L1; 1.
DR  PRINTS; PR00865; HPVcapsidL1.
DR  ProDom; PD000544; PV_capsid_L1; 1.
FT  NON_TER 1
FT  NON_TER 494
SQ  SEQUENCE 494 AA; 55021 MW; B2D7AC39D176C3C5 CRC64;

Query Match          100.0%; Score 49; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RT populations characterized by nucleotide sequence analysis of the E6,
RT L2, and L1 coding segments.";
RT J. Virol. 69:7743-7753 (1995).

RESULT 16

```

O81012          PRELIMINARY;      PRT;      494 AA.
ID  O81012;
AC  O81012;
DT  01-NOV-1996 (TREMblrel. 01, Created)
DT  01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT  01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE  Late major capsid protein L1 (Fragment).
GN  L1.
OS  Human papillomavirus.
OC  Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC  Papillomavirus.
OX  NCBI_TaxID=10566;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=OR_9237;
RX  MEDLINE=96079021; PubMed=7494284;
RA  Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA  Jenison S.A.;
RT  "Human papillomavirus type 16 variant lineages in United States
RT  populations characterized by nucleotide sequence analysis of the E6,
RT  L2, and L1 coding segments.";
RT  J. Virol. 69:7743-7753 (1995).
RN  (2)
RN  SEQUENCE FROM N.A.
RP  STRAIN=OR_9237;
RC  Farmer A.D.;
RA  Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR  EMBL; U34163; AAA91716.1; -.
DR  GO; GO:0019028; C:Viral capsid; IEA.
DR  GO; GO:0005198; F:Structural molecule activity; IEA.
DR  InterPro; IPR002210; PV_capsid_L1.
DR  InterPro; IPR008975; Viral_cap_coat.
DR  Pfam; PF00500; late_protein_L1; 1.
DR  PRINTS; PR00865; HPVcapsidL1.
DR  ProDom; PD000544; PV_capsid_L1; 1.
FT  NON_TER 1
FT  NON_TER 494
SQ  SEQUENCE 494 AA; 55094 MW; ECB31B09E8E8E8E3 CRC64;

Query Match          100.0%; Score 49; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 AGVDNRECI 9
    |||||
    |||||
    |||||
Db  130 AGVDNRECI 138

```

RESULT 17

```

O81017          PRELIMINARY;      PRT;      494 AA.
ID  O81017;
AC  O81017;
DT  01-NOV-1996 (TREMblrel. 01, Created)
DT  01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT  01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE  Late major capsid protein L1 (Fragment).
GN  L1.
OS  Human papillomavirus.
OC  Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC  Papillomavirus.
OX  NCBI_TaxID=10566;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=OR_9237;
RX  MEDLINE=96079021; PubMed=7494284;
RA  Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA  Jenison S.A.;
RT  "Human papillomavirus type 16 variant lineages in United States
RT  populations characterized by nucleotide sequence analysis of the E6,
RT  L2, and L1 coding segments.";
RT  J. Virol. 69:7743-7753 (1995).
RN  (2)
RN  SEQUENCE FROM N.A.
RP  STRAIN=OR_9237;
RC  Farmer A.D.;
RA  Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR  EMBL; U34163; AAA91716.1; -.
DR  GO; GO:0019028; C:Viral capsid; IEA.
DR  GO; GO:0005198; F:Structural molecule activity; IEA.
DR  InterPro; IPR002210; PV_capsid_L1.
DR  InterPro; IPR008975; Viral_cap_coat.
DR  Pfam; PF00500; late_protein_L1; 1.
DR  PRINTS; PR00865; HPVcapsidL1.
DR  ProDom; PD000544; PV_capsid_L1; 1.
FT  NON_TER 1
FT  NON_TER 494
SQ  SEQUENCE 494 AA; 55021 MW; B2D7AC39D176C3C5 CRC64;

Query Match          100.0%; Score 49; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Oy      1 AGVDNRECT 9
Db      130 AGVDNRECT 138

RESULT 17
081001 PRELIMINARY; PRT; 494 AA.
AC 081001.
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 1905;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
RT populations characterized by nucleotide sequence analysis of the E6,
RT L2, and L1 coding segments."
RT J. Virol. 69:7743-7753(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 1905;
RA Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U34172; AAA91719.1; -
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT 494 494
SQ SEQUENCE 494 AA; 55094 MW; ECB31B09B8EF4E3 CRC64;

Query Match 100.0%; Score 49; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 AGVDNRECT 9
Db      130 AGVDNRECT 138

RESULT 18
090204 PRELIMINARY; PRT; 494 AA.
AC 090204.
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96079021; PubMed=7494284;
RX Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
```

```
RT populations characterized by nucleotide sequence analysis of the E6,
RT L2, and L1 coding segments."
RT J. Virol. 69:7743-7753(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U34173; AAA91720.1; -
DR EMBL; U34175; AAA91722.1; -
DR EMBL; U34178; AAA91725.1; -
DR EMBL; U34179; AAA91726.1; -
DR EMBL; U34168; AAA91715.1; -
DR EMBL; U34171; AAA91718.1; -
DR EMBL; U34180; AAA91727.1; -
DR EMBL; U34191; AAA91738.1; -
DR EMBL; U34192; AAA91739.1; -
DR EMBL; U34182; AAA91729.1; -
DR EMBL; U34185; AAA91732.1; -
DR EMBL; U34186; AAA91733.1; -
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT 494 494
SQ SEQUENCE 494 AA; 55021 MW; A737AC39CB6CC3DF CRC64;

Query Match 100.0%; Score 49; DB 12; Length 494;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 AGVDNRECT 9
Db      130 AGVDNRECT 138

RESULT 19
081000 PRELIMINARY; PRT; 494 AA.
AC 081000.
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Late major capsid protein L1 (Fragment).
GN L1.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NM T529;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA Jenison S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
RT populations characterized by nucleotide sequence analysis of the E6,
RT L2, and L1 coding segments."
RT J. Virol. 69:7743-7753(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NM T529;
RA Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U34170; AAA91717.1; -
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
```

```

DR PRINTS: PR00865; HPVcapsid.L1.
DR Prodrom; PD000544; PV_capsid_L1; 1.
RA NON_TER 1
FT NON_TER 494
SQ SEQUENCE 494 AA; 55063 MW; 735F7C2B86CACC58F CRC64;

Query Match
Best Local Similarity 100.0%; Score 49; DB 12; Length 494;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
Db 130 AGVDNRECI 138

RESULT 20
ID 012343 PRELIMINARY; PRT; 495 AA.
AC 012343;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DE Late major capsid protein L1 (Fragment).
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCB1_Taxid=10581;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=PC-8;
RX MEDLINE=97437474; PubMed=9292007;
RA Torresello M.L., Buonaguro F.M., Meglio A., Buonaguro L.,
RA Beth-Giraldo E., Giraldo G.;
RT "Sequence variations and viral genomic state of human papillomavirus
RT type 16 in penile carcinomas from Ugandan patients.";
RL J. Gen. Virol. 78:2199-2208(1997).
DR EMBL; AF003031; AAB70748.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsid.L1.
DR Prodrom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 495
SQ SEQUENCE 495 AA; 55133 MW; 427D4FC2B42E1EDD CRC64;

Query Match
Best Local Similarity 100.0%; Score 49; DB 12; Length 495;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
Db 133 AGVDNRECI 141

RESULT 21
Q9WHS7 PRELIMINARY; PRT; 505 AA.
ID 09WHS7;
AC 09WHS7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE Capsid protein.
DE Capsid protein.
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCB1_Taxid=10581;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=GU2;

```

```

RX MEDLINE=99252179; PubMed=10233949;
RA White W.I., Wilson S.D., Palmer-Hill F.J., Woods R.M., Gim S.J.,
RA Hewitt L.A., Goldman D.M., Burke S.J., Jensen A.B., Koenig S.,
RA Suzich J.A.;
RT "Characterization of a major neutralizing epitope on human
RT papillomavirus type 16 L1.";
RL J. Virol. 73:4882-4889(1999).
DR EMBL; AF134178; AAD22113.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsid.L1.
DR Prodrom; PD000544; PV_capsid_L1; 1.
SQ SEQUENCE 505 AA; 56245 MW; 37E59806BF61A5F CRC64;

Query Match
Best Local Similarity 100.0%; Score 49; DB 12; Length 505;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
Db 139 AGVDNRECI 147

RESULT 22
Q9WLOS PRELIMINARY; PRT; 505 AA.
ID 09WLOS;
AC 09WLOS;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE Late major capsid protein.
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCB1_Taxid=10581;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Phil;
RA Touze A., El-Mehdaoui S., Munoz N., Mouglin C., Coursaget P.;
RT "Papillomavirus type 16 full length sequence of L1 gene.";
RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF043287; AAC09293.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV_capsid_L1.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsid.L1.
DR Prodrom; PD000544; PV_capsid_L1; 1.
SQ SEQUENCE 505 AA; 56332 MW; C348AC9308D795ED CRC64;

Query Match
Best Local Similarity 100.0%; Score 49; DB 12; Length 505;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
Db 139 AGVDNRECI 147

RESULT 23
Q9WHS6 PRELIMINARY; PRT; 505 AA.
ID 09WHS6;
AC 09WHS6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE Capsid protein.
GN L1.

```

OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GU1;
RX MEDLINE=98252179; PubMed=10233949;
RA White W.I., Wilson S.D., Palmer-Hill F.J., Woods R.M., Chim S.J.,
Hewitt L.A., Goldman D.M., Burke S.J., Jensen A.B., Koenig S.,
Ruzich J.A.;
RT "Characterization of a major neutralizing epitope on human
papillomavirus type 16 L1";
RL J. Virol. 73:4882-4889(1999).
DR EMBL; AF134177; AAD22112.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR InterPro; IPR002210; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV capsid L1.
DR InterPro; IPR008975; Viral cap_coat.
DR Pfam; PF00500; late protein L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR PRODOM; PD000544; PV capsid L1; 1.
SQ SEQUENCE 505 AA; 56260 MW; 43B61B624BED85 CRC64;
QY Query Match 100.0%; Score 49; DB 12; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 139 AGVDNRECT 147
RESULT 24
ID 081015 PRELIMINARY; PRT; 505 AA.
AC 081015;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE L1 capsid protein.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1194;
RX MEDLINE=96130211; PubMed=8553573;
RA Icenogle J.P., Ciancy K.A., Lin S.Y.;
RT "Sequence variation in the capsid protein genes of human
papillomavirus type 16 and type 31.";
RL Virology 214:664-669(1995).
DR EMBL; U37217; AAA92892.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV capsid L1.
DR InterPro; IPR008975; Viral cap_coat.
DR Pfam; PF00500; late protein L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR PRODOM; PD000544; PV capsid L1; 1.
SQ SEQUENCE 505 AA; 56344 MW; 52D2D06922836F68 CRC64;
QY Query Match 100.0%; Score 49; DB 12; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 139 AGVDNRECT 147
RESULT 25

Q9WPH4
ID Q9WPH4 PRELIMINARY; PRT; 505 AA.
AC Q9WPH4;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang L.L., Song G.X.;
RT "The L1 Variant of Human Papillomavirus Type 16 in Condyloma
Acuminatum Tissue from one Chinese Woman";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF140365; AAD38523.1; -;
DR PDB; 1DZL; 25-AUG-00.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV capsid L1.
DR InterPro; IPR008975; Viral cap_coat.
DR Pfam; PF00500; late protein L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR PRODOM; PD000544; PV capsid L1; 1.
SQ SEQUENCE 505 AA; 56179 MW; 7526CECAC6B0E4 CRC64;
QY Query Match 100.0%; Score 49; DB 12; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 139 AGVDNRECT 147
RESULT 26
ID 092282 PRELIMINARY; PRT; 505 AA.
AC 092282;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Late major capsid protein.
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang L.L., Song G.X., Si J.Y.;
RT "Human papillomavirus type 16 late major capsid gene (L1).";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF084952; AAC61736.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV capsid L1.
DR InterPro; IPR008975; Viral cap_coat.
DR Pfam; PF00500; late protein L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR PRODOM; PD000544; PV capsid L1; 1.
SQ SEQUENCE 505 AA; 56308 MW; 852BA3B3C09FD5EB CRC64;
QY Query Match 100.0%; Score 49; DB 12; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 139 AGVDNRECT 147

Query Match
Best Local Similarity 100.0%; Score 49; DB 12; Length 531;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGVDNRECT 9
Db 165 AGVDNRECT 173

RESULT 30
Q8BSN9 PRELIMINARY; PRT; 531 AA.
AC Q8BSN9;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative major capsid protein L1.
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RA Teraï M., Ma Z., Burk R.D.;
RT "Cloning and Sequencing of Non-European Human Papillomavirus (HPV)
RT Variants Complete Genomes from Patients with Cervical Cancer by an
RT Overlapping PCR Method.";
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF472508; AA015704.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV capsid L1.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR PRODOM; PD000544; PV capsid L1; 1.
DR SEQUENCE 531 AA; 59548 MW; 320626E6A4F6EFA3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 49; DB 12; Length 531;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGVDNRECT 9
Db 165 AGVDNRECT 173

RESULT 30
Q8BSN9 PRELIMINARY; PRT; 531 AA.
AC Q8BSN9;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative major capsid protein L1.
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RA Teraï M., Ma Z., Burk R.D.;
RT "Cloning and Sequencing of Non-European Human Papillomavirus (HPV)
RT Variants Complete Genomes from Patients with Cervical Cancer by an
RT Overlapping PCR Method.";
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF472508; AA015704.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV capsid L1.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR PRODOM; PD000544; PV capsid L1; 1.
DR SEQUENCE 531 AA; 59548 MW; 320626E6A4F6EFA3 CRC64;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGVDNRECT 9
Db 165 AGVDNRECT 173

RESULT 31
Q8BSN5 PRELIMINARY; PRT; 531 AA.
AC Q8BSN5;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative major capsid protein L1.
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RA Teraï M., Ma Z., Burk R.D.;
RT "Cloning and Sequencing of Non-European Human Papillomavirus (HPV)
RT Variants Complete Genomes from Patients with Cervical Cancer by an
RT Overlapping PCR Method.";
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF472509; AA015712.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV capsid L1.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR PRODOM; PD000544; PV capsid L1; 1.
DR SEQUENCE 531 AA; 59541 MW; F1C433C8ADA313AA CRC64;

Query Match
Best Local Similarity 100.0%; Score 49; DB 12; Length 531;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGVDNRECT 9
Db 165 AGVDNRECT 173

RESULT 32
Q8OML8 PRELIMINARY; PRT; 531 AA.
AC Q8OML8;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE L1.
GN L1.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RA Teraï M., Burk R.D.;
RT "Human papillomavirus type 16 Asian-American variant.";
RT Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF402678; AA085415.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002210; PV capsid L1.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR PRODOM; PD000544; PV capsid L1; 1.
DR SEQUENCE 531 AA; 59517 MW; 72E7F90A83D48186 CRC64;

Query Match 100.0%; Score 49; DB 12; Length 531;
 Best Local Similarity 100.0%; Pred. No. 0.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
 |||||
 DB 165 AGVDNRECI 173

RESULT 33

Q81017 PRELIMINARY; PRT; 504 AA.

AC 081017;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE L1 capsid protein.
 OS Human papillomavirus type 31.
 OC Viruses; deDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_Taxid=10585;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=MS2;
 RX Icenogle J.P., Clancy K.A., Lin S.Y.;
 RT "Sequence variation in the capsid protein genes of human
 papillomavirus type 16 and type 31."
 RL Virology 214:664-669 (1995).
 DR EMBL; U37410; AAA92894.1; -
 DR GO; GO:0019028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR002210; PV_Capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR Prodom; PD000544; PV capsid L1; 1.
 SO SEQUENCE 504 AA; 56335 MW; 0A3970E631E2FA7 CRC64;

Query Match 83.7%; Score 41; DB 12; Length 504;
 Best Local Similarity 87.5%; Pred. No. 4.8;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVDNRECI 9
 |||||
 DB 141 GTDNRECI 148

RESULT 34

Q90730 PRELIMINARY; PRT; 529 AA.

AC 090730;
 DT 01-NOV-1998 (TRENBLREL. 08, Created)
 DT 01-NOV-1998 (TRENBLREL. 08, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE ORF L1.
 OS Human papillomavirus type 67.
 OC Viruses; deDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_Taxid=37120;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Kiril Y., Iwamoto S., Matsukura T.;
 RT "Complete DNA sequence of human Papillomavirus type 67 isolated from a
 vaginal intraepithelial neoplasia."
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9073695; PubMed=9857984;
 RA Kiril Y., Matsukura T.;
 RT "Nucleotide sequence and phylogenetic classification of
 humanpapillomavirus type 67.";

RL Virus Genes 17:117-121 (1998).

DR EMBL; D21208; BAA28859.1; -
 DR GO; GO:0019028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR002210; PV_Capsid_L1.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00500; late_protein_L1; 1.
 DR PRINTS; PR00865; HPVcapsidL1.
 DR Prodom; PD000544; PV capsid L1; 1.
 SO SEQUENCE 529 AA; 59425 MW; 9750C245F1BC6C07 CRC64;

Query Match 79.6%; Score 39; DB 12; Length 529;
 Best Local Similarity 75.0%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVDNRECI 9
 |||||
 DB 170 GTDNRECI 177

RESULT 35

Q7UKH0 PRELIMINARY; PRT; 310 AA.

AC 07UKH0;
 DT 01-OCT-2003 (TRENBLREL. 25, Created)
 DT 01-OCT-2003 (TRENBLREL. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE Hypothetical protein.
 OS Rhodospirillum rubrum.
 GN Rbl0648.
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Pirellula.
 OX NCBI_Taxid=117;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=1;
 RX MEDLINE=22735913; PubMed=12835416;
 RA Glockner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
 RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,
 RA Schlesner H., Amann R., Reinhardt R.;
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.
 strain 1."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
 DR EMBL; BX294152; CAD76894.1; -
 KM Hypothetical protein; Complete proteome.
 SO SEQUENCE 310 AA; 34620 MW; B5CA0784C740A8AA CRC64;

Query Match 73.5%; Score 36; DB 16; Length 310;
 Best Local Similarity 66.7%; Pred. No. 33;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
 |||||
 DB 69 AGVDNRECI 77

RESULT 36

Q9AMT8 PRELIMINARY; PRT; 445 AA.

AC 09AMT8;
 DT 01-JUN-2001 (TRENBLREL. 17, Created)
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
 DE ID872 (EC 6.3.1.2) (Glutamine synthetase) (glutamate--ammonia ligase)
 DE (B1r2084 protein).
 GN ID872 OR B1r2084.
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 OX NCBI_Taxid=375;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=110epc4;

RX MEDLINE=21101824; PubMed=1157954;
 RA Gottfert M., Rothliberger S., Kundig C., Beck C., Marty R.,
 RA Hennecke H.,
 RT "Potential symbiosis-specific genes uncovered by sequencing a 410-kb
 RT DNA region of the *Bradyrhizobium japonicum* chromosome.";
 RL J. Bacteriol. 183:1405-1412(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-USDA 110;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Kaneo T., Nakamura Y., Sato S., Minamisawa K., Uchiuni T.,
 RA Sasamoto S., Matsubae A., Ideasa K., Irituchi M., Kawashima K.,
 RA Kohara M., Matsunoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
 RA Tabata S.;
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
 RT *Bradyrhizobium japonicum* USDA110.";
 RL DNA Res. 9:189-197(2002).
 CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) = ADP + PHOSPHATE +
 CC L-GLUTAMINE.
 CC -1- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
 DR EMBL; AF322013; AAC61059.1; -;
 DR EMBL; AP005942; BAC47349.1; -;
 DR GO; GO:0004356; F:glutamate- ammonia ligase activity; IEA.
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR GO; GO:0009399; P:nitrogen fixation; IEA.
 DR InterPro; IPR008146; Gln_synth_C.
 DR Pfam; PF00120; gln-synth_1.
 DR Prodom; PD001057; Gln_synth_C_1.
 DR LIGASE; Complete proteome.
 KW SEQUENCE 445 AA; 48649 MW; AE1A8966A279388 CRC64;

Query Match 73.5%; Score 36; DB 16; Length 445;
 Best Local Similarity 85.7%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVDNREC 8
 Db 318 GVDNREC 324

RESULT 37
 O93879 PRELIMINARY; PRT; 474 AA.
 ID O93879
 AC O93879
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE K09E9.3 protein.
 GN K09E9.3.
 OS *Caenorhabditis elegans*.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; *Caenorhabditis*.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Swindburne J., Ainscough R.;
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C.elegans*: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z79602; CAB01890.1; -;
 DR PIR; T23547; T23547.
 DR WormPep; K09E9.3; CE12014.
 SQ SEQUENCE 474 AA; 54474 MW; A4710D09946143AA CRC64;

Query Match 73.5%; Score 36; DB 5; Length 474;
 Best Local Similarity 66.7%; Pred. No. 51;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 AGVDNREC 9
 Db 275 AGVDNREC 283

RESULT 38
 O95R86 PRELIMINARY; PRT; 539 AA.
 ID O95R86
 AC O95R86
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE SD01935D.
 GN HMS-BEAGLE\POL OR CG2485 OR CG17399.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY061571; AAL29119.1; -;
 DR FlyBase; FBgn0044441; HMS-Beagle\POL.
 SQ SEQUENCE 539 AA; 61775 MW; 8698C687AF079D0B CRC64;

Query Match 73.5%; Score 36; DB 5; Length 539;
 Best Local Similarity 55.6%; Pred. No. 58;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 AGVDNREC 9
 Db 333 SMDNREC 341

RESULT 39
 O9M924 PRELIMINARY; PRT; 1039 AA.
 ID O9M924
 AC O9M924
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative N-arginine dipeptide convertase.
 GN P4H5.4
 OS *Arabidopsis thaliana* (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; *Arabidopsis*.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
 RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
 RA Gonzalez A., Kremetska I., Kim C., Lenz C., Li J., Liu S.,
 RA Luvos S., Schwartz J., Shinn P., Toriumi M., Vysotskaya V.S.,
 RA Walker M., Yu G., Becker J., Theologis A., Davis R.W.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC011001; AAF63132.1; -;
 DR PIR; G86203; G86203.
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001431; peptidase_M16.
 DR InterPro; IPR007863; peptidase_M16_C.
 DR Pfam; PF00675; peptidase_M16_1.
 DR Pfam; PF05193; peptidase_M16_C_2.
 DR PROSITE; PS00143; INSULINASE; 1.
 SQ SEQUENCE 1039 AA; 119123 MW; 077DE73D1308A8D CRC64;

Query Match 73.5%; Score 36; DB 10; Length 1039;
 Best Local Similarity 87.5%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVDNRECI 9
 DB 271 GVDLRECI 278

RESULT 40

Q82640 PRELIMINARY; PRT; 243 AA.
 AC Q82640;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative bacteriophage protein.
 GN STY2021 OR T1875.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18:
 RX MEDLINE=21534947; PubMed=11677608;
 RA ParKhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jørgensen K.,
 RA Krogh A., Larsen T.S., Leather S., Mouton S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RA "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 DR EMBL; AL627272; CAD05564.1; -;
 DR EMBL; AE016840; AA069492.1; -;
 DR InterPro: IPR008941; TPR-like.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 243 AA; 27009 MW; CF43169FB3104D6 CRC64;

Query Match 71.4%; Score 35; DB 16; Length 243;
 Best Local Similarity 55.6%; Pred. No. 42;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AGVDNRECI 9
 DB 227 AGLENSCV 235

Search completed: March 3, 2004, 10:24:47
 Job time : 37.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 10:23:24 ; Search time 27 Seconds
(without alignments)
70.384 Million cell updates/sec

Title: US-09-980-064-2
Perfect score: 49
Sequence: 1 AGVDNRECT 9

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 segs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	505	9 US-09-162-904A-3	Sequence 3, Appli
2	49	100.0	505	9 US-09-162-904A-4	Sequence 4, Appli
3	49	100.0	505	9 US-09-824-017-2	Sequence 2, Appli
4	49	100.0	505	10 US-09-986-118A-2	Sequence 2, Appli
5	49	100.0	505	15 US-10-367-095-6	Sequence 1, Appli
6	49	100.0	531	10 US-09-991-053-1	Sequence 6, Appli
7	49	100.0	531	14 US-10-301-260A-1	Sequence 1, Appli
8	42	85.7	502	14 US-10-224-999A-3469	Sequence 3469, Ap
9	35	71.4	126	8 US-08-450-842-57	Sequence 57, Appli
10	34	69.4	26	14 US-10-094-401-253	Sequence 23, App
11	34	69.4	26	15 US-10-462-262-217	Sequence 217, App
12	34	69.4	169	14 US-10-177-390-28	Sequence 28, Appli
13	34	69.4	874	15 US-10-294-006-34	Sequence 34, Appli
14	34	69.4	2444	9 US-09-944-849-2	Sequence 2, Appli
15	34	69.4	2469	15 US-10-190-115-2	Sequence 2, Appli

16	34	69.4	2469	15 US-10-369-072-2	Sequence 2, Appli
17	34	69.4	2471	15 US-10-190-115-27	Sequence 27, Appli
18	34	69.4	2471	15 US-10-369-072-27	Sequence 27, Appli
19	34	69.4	2531	15 US-10-190-115-29	Sequence 29, Appli
20	34	69.4	2531	15 US-10-369-072-29	Sequence 29, Appli
21	34	69.4	2556	15 US-10-294-006-12	Sequence 12, Appli
22	33	67.3	316	15 US-10-369-493-13372	Sequence 13372, A
23	33	67.3	429	14 US-10-156-461-14060	Sequence 14060, A
24	33	67.3	984	14 US-10-354-358-102	Sequence 102, App
25	33	67.3	984	15 US-10-116-575-147	Sequence 147, App
26	32	65.3	135	9 US-09-901-436A-8	Sequence 8, Appli
27	32	65.3	135	13 US-10-032-366-3	Sequence 3, Appli
28	32	65.3	135	13 US-10-032-366-4	Sequence 4, Appli
29	32	65.3	135	14 US-10-171-311-200	Sequence 200, App
30	32	65.3	219	15 US-10-264-049-3046	Sequence 3046, App
31	32	65.3	221	15 US-10-369-493-12041	Sequence 12041, A
32	32	65.3	308	15 US-10-369-493-2249	Sequence 2249, Ap
33	32	65.3	442	9 US-09-989-545-17	Sequence 17, Appli
34	32	65.3	457	15 US-10-446-520-14	Sequence 14, Appli
35	32	65.3	787	15 US-10-369-493-9643	Sequence 9643, Ap
36	32	65.3	1115	10 US-09-985-536A-10	Sequence 10, Appli
37	32	65.3	1115	14 US-10-222-668-3	Sequence 3, Appli
38	32	65.3	1285	14 US-10-205-823-273	Sequence 273, App
39	32	65.3	1285	14 US-10-177-293-317	Sequence 317, App
40	31	63.3	92	9 US-08-864-761-48480	Sequence 48480, A
41	31	63.3	198	14 US-10-319-799-40	Sequence 40, Appli
42	31	63.3	203	15 US-10-369-493-8374	Sequence 8374, Ap
43	31	63.3	265	9 US-09-870-759-108	Sequence 108, App
44	31	63.3	265	10 US-09-751-708A-108	Sequence 108, App
45	31	63.3	419	15 US-10-369-493-19381	Sequence 19381, A

ALIGNMENTS

RESULT 1
US-09-162-904A-3
Sequence 3, Application US/09162904A
Patent No. US20020168372A1
GENERAL INFORMATION:
APPLICANT: Durel, Mattheia
TITLE OF INVENTION: OF A DNA SEQUENCE ENCODING A PAPILLOMAVIRUS L1 PROTEIN CAPABLE
FILE REFERENCE: 8484-068-999
CURRENT FILING DATE: 1998-09-29
PRIOR FILING DATE: 1997-06-27
PRIOR APPLICATION NUMBER: 08/641,570
PRIOR FILING DATE: 1996-05-01
PRIOR APPLICATION NUMBER: 08/902,528
PRIOR FILING DATE: 1993-07-16
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 505
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Amino acid sequence of L1 gen
US-09-162-904A-3

Query Match 100.0%; Score 49; DB 9; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGVDNRECT 9
Db 139 AGVDNRECT 147

RESULT 2
US-09-162-904A-4
Sequence 4, Application US/09162904A
Patent No. US20020168372A1
GENERAL INFORMATION:
APPLICANT: Dursch, Matthias
APPLICANT: Gissmann, Inez
TITLE OF INVENTION: A DNA SEQUENCE ENCODING A PAPILLOMAVIRUS L1 PROTEIN CAPABLE
OF EFFICIENTLY FORMING VIRUS-LIKE PARTICLES
FILE REFERENCE: 8484-068-999
CURRENT APPLICATION NUMBER: US/09/162,904A
CURRENT FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 08/884,168
PRIOR FILING DATE: 1997-06-27
PRIOR APPLICATION NUMBER: 08/641,570
PRIOR FILING DATE: 1996-05-01
PRIOR APPLICATION NUMBER: 08/902,528
PRIOR FILING DATE: 1993-07-16
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin version 3.1
SEQ ID NO 4
LENGTH: 505
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Amino acid sequence of L1 ger
US-09-162-904A-4

Query Match 100.0%; Score 49; DB 9; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
DB 139 AGVDNRECI 147

RESULT 3
US-09-824-017-2
Sequence 2, Application US/09824017
Publication No. US20020197668A1
GENERAL INFORMATION:
APPLICANT: BURGER, Alexander
APPLICANT: HALLER, Michael
TITLE OF INVENTION: PAPILLOMA VIRUS CAPSOMERE VACCINE
FORMULATIONS AND METHODS OF USE
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/824,017
FILING DATE: 03-Apr-2001
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/026,896
FILING DATE: 1998-02-20
ATTORNEY/AGENT INFORMATION:
NAME: Sandercock, Colin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 37067/102
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300

TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-824-017-2

Query Match 100.0%; Score 49; DB 9; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
DB 139 AGVDNRECI 147

RESULT 4
US-09-986-118A-2
Sequence 2, Application US/09986118A
Publication No. US20030021806A1
GENERAL INFORMATION:
APPLICANT: BURGER, Alexander
APPLICANT: HALLER, Michael
TITLE OF INVENTION: PAPILLOMA VIRUS CAPSOMERE VACCINE
FORMULATIONS AND METHODS OF USE
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/986,118A
FILING DATE: 07-No. US20030021806A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/026,896
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Sandercock, Colin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 37067/102
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-986-118A-2

Query Match 100.0%; Score 49; DB 10; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
DB 139 AGVDNRECI 147

```
RESULT 5
US-10-367-095-6
; Sequence 6, Application US/10367095
; Publication No. US20030228696A1
; GENERAL INFORMATION:
; APPLICANT: Robin A. Robinson
; TITLE OF INVENTION: No. US20030228696A1 Insect Cell Line
; FILE REFERENCE: 44149-1US1
; CURRENT APPLICATION NUMBER: US/10/367,095
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 60/356,119
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,161
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,118
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,133
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,157
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,156
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,123
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,113
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,154
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,135
; PRIOR FILING DATE: 2002-02-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HPV-16 L1 protein
US-10-367-095-6

Query Match          100.0%; Score 49; DB 15; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECT 9
Db 139 AGVDNRECT 147

RESULT 6
US-09-991-053-1
; Sequence 1, Application US/09991053
; Publication No. US20030003532A1
; GENERAL INFORMATION:
; APPLICANT: Harrison, Stephen
; APPLICANT: Chen, Xiaojiang
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
; FILE REFERENCE: Harvard/Harrison 12687/1120
; CURRENT APPLICATION NUMBER: US/09/991,053
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-18
; PRIOR APPLICATION NUMBER: 60/125208
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/148544
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 531
; TYPE: PRT
```

```
; ORGANISM: Human papillomavirus type 16
US-09-991-053-1

Query Match          100.0%; Score 49; DB 10; Length 531;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECT 9
Db 165 AGVDNRECT 173

RESULT 7
US-10-301-260A-1
; Sequence 1, Application US/10301260A
; Publication No. US20030118609A1
; GENERAL INFORMATION:
; APPLICANT: Harrison, Stephen
; APPLICANT: Chen, Xiaojiang
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
; FILE REFERENCE: Harvard/Harrison 12687/1123
; CURRENT APPLICATION NUMBER: US/10/301,260A
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: 09/520,822
; PRIOR FILING DATE: 2000-03-18
; PRIOR APPLICATION NUMBER: 60/125208
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/148544
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-301-260A-1

Query Match          100.0%; Score 49; DB 14; Length 531;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECT 9
Db 165 AGVDNRECT 173

RESULT 8
US-10-224-999A-3469
; Sequence 3469, Application US/10224999A
; Publication No. US20030171318A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Composition and Method for Treating Viral Infection
; FILE REFERENCE: 5004.01
; CURRENT APPLICATION NUMBER: US/10/224,999A
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/313,695
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 3484
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3469
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Human papillomavirus type 35
US-10-224-999A-3469

Query Match          85.7%; Score 42; DB 14; Length 502;
Best Local Similarity 77.8%; Pred. No. 4.9;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 AGVDNRECT 9
DB 139 SGTDRRECT 147

RESULT 9

US-08-450-842-57
; Sequence 57, Application US/08450842
; Publication No. US20020045576A1
; GENERAL INFORMATION:
; APPLICANT: GENESENTECH, INC.
; APPLICANT: ROSENTHAL, ARNON
; TITLE OF INVENTION: NOVEL NEUTROTROPIC FACTOR
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,842
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/426419
; FILING DATE: 19-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/030013
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/648482
; FILING DATE: 31-JAN
; APPLICATION DATA:
; APPLICATION NUMBER: 07/587707
; FILING DATE: 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: 666P2CID3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8674
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-450-842-57

Query Match 71.4%; Score 35; DB 8; Length 126;
Best Local Similarity 85.7%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVDNRECT 8
DB 80 GVDNRECT 86

RESULT 10
US-10-094-401-253
; Sequence 253, Application US/10094401
; Publication No. US20030069395A1
; GENERAL INFORMATION:
; APPLICANT: DYAX CORP.

; APPLICANT: Sato, Aaron K.
; APPLICANT: Ley, Arthur C.
; APPLICANT: Cohen, Edward H.
; TITLE OF INVENTION: SERUM ALBUMIN BINDING MOETIES
; FILE REFERENCE: DYX-026.2 PCT; DYX-026.2 US
; CURRENT APPLICATION NUMBER: US/10/094,401
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: 60/331,352
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/292,975
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 253
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: albumin binding peptide
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: ACETYLATION
; NAME/KEY: MOD RES
; LOCATION: (26)..(26)
; OTHER INFORMATION: AMIDATION
US-10-094-401-253

Query Match 69.4%; Score 34; DB 14; Length 26;
Best Local Similarity 83.3%; Pred. No. 8.2;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DNRECT 9
DB 2 DNRECT 7

RESULT 11
US-10-462-262-217
; Sequence 217, Application US/10462262
; Publication No. US20040009534A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Aaron K.
; APPLICANT: Dawson, Bruce M.
; TITLE OF INVENTION: PROTEIN ANALYSIS
; FILE REFERENCE: 10280-052001
; CURRENT APPLICATION NUMBER: US/10/462,262
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/388,642
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 217
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: serum albumin-binding agent
US-10-462-262-217

Query Match 69.4%; Score 34; DB 15; Length 26;
Best Local Similarity 83.3%; Pred. No. 8.2;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DNRECT 9
DB 2 DNRECT 7

RESULT 12
US-10-177-390-28
; Sequence 28, Application US/10177390
; Publication No. US20030143743A1
; GENERAL INFORMATION:


```
;; APPLICANT: Schuler, Gerold
;; APPLICANT: N.V. Antwerp Innovatiecentrum
;; TITLE OF INVENTION: Improved Transfection of Eucaryotic Cells with Linear
;; FILE REFERENCE: Polynucleotides by Electroporation
;; CURRENT APPLICATION NUMBER: US/10/177,390
;; CURRENT FILING DATE: 2002-06-20
;; NUMBER OF SEQ ID NOS: 34
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 28
;; LENGTH: 169
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-177-390-28

Query Match          69.4%; Score 34; DB 14; Length 169;
Best Local Similarity 57.1%; Pred. No. 59;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 VDNRECI 9
Db      101 IDNRQCV 107

RESULT 13
US-10-294-006-34
;; Sequence 34, Application US/10294006
;; Publication No. US20040013657A1
;; GENERAL INFORMATION:
;; APPLICANT: Tang, Y. Tom
;; APPLICANT: Zhou, Ping
;; APPLICANT: Goodrich, Kyle
;; APPLICANT: Asundi, Vinod
;; APPLICANT: Yang, Yonghong
;; APPLICANT: Zhang, Jie
;; APPLICANT: Wehrman, Tom
;; APPLICANT: Drmanac, Radoje T.
;; TITLE OF INVENTION: No. US20040013657A1el Nucleic Acids and
;; FILE REFERENCE: Polypeptides
;; CURRENT APPLICATION NUMBER: US/10/294,006
;; CURRENT FILING DATE: 2002-11-12
;; PRIOR APPLICATION NUMBER: 09/815,925
;; PRIOR FILING DATE: 2001-03-22
;; NUMBER OF SEQ ID NOS: 44
;; SOFTWARE: pt_FL_genes Version 2.0
;; SEQ ID NO 34
;; LENGTH: 874
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc.feature
;; LOCATION: (1) ..(874)
;; OTHER INFORMATION: Xaa = any amino acid or symbol as shown in the table 8 as set for
US-10-294-006-34

Query Match          69.4%; Score 34; DB 15; Length 874;
Best Local Similarity 57.1%; Pred. No. 3,4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 VDNRECI 9
Db      812 IDNRQCV 818

RESULT 14
US-09-944-849-2
;; Sequence 2, Application US/09944849
;; Patent No. US20020151487A1
;; GENERAL INFORMATION:
;; APPLICANT: Nickoloff, Brian
;; APPLICANT: Miele, Lucio
;; TITLE OF INVENTION: METHOD AND REAGENTS FOR EPITHELIAL BARRIER FORMATION AND TREATMEN
```

```
;; TITLE OF INVENTION: MALIGNANT AND BENIGN SKIN DISORDERS BY MODULATING THE NOTCH PATH
;; FILE REFERENCE: 212583
;; CURRENT APPLICATION NUMBER: US/09/944,849
;; CURRENT FILING DATE: 2001-08-31
;; PRIOR APPLICATION NUMBER: US 60/229,614
;; PRIOR FILING DATE: 2000-08-31
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 2
;; LENGTH: 2444
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; NAME/KEY: misc.feature
;; LOCATION: (891)..(891)
;; OTHER INFORMATION: The 'Xaa' at location 891 stands for Gly, or Ala.
;; NAME/KEY: misc.feature
;; LOCATION: (1763)..(1763)
;; OTHER INFORMATION: The 'Xaa' at location 1763 stands for Gln, Arg, Pro, or Leu.
;; NAME/KEY: misc.feature
;; LOCATION: (1787)..(1787)
;; OTHER INFORMATION: The 'Xaa' at location 1787 stands for Thr, Ala, Pro, or Ser.
;; OTHER INFORMATION: Constitutively Active No. US20020151487A1ch-1
US-09-944-849-2

Query Match          69.4%; Score 34; DB 9; Length 2444;
Best Local Similarity 57.1%; Pred. No. 1e+03;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 VDNRECI 9
Db      1681 IDNRQCV 1687
```

```
RESULT 15
US-10-190-115-2
;; Sequence 2, Application US/10190115
;; Publication No. US20030207394A1
;; GENERAL INFORMATION:
;; APPLICANT: Alsbrook, John P. II
;; APPLICANT: Boldog, Ferenc L.
;; APPLICANT: Burgess, Catherine E.
;; APPLICANT: Casman, Stacie J.
;; APPLICANT: Groese, William M.
;; APPLICANT: Gusev, Vladimir Y.
;; APPLICANT: Ji, Weizhen
;; APPLICANT: Lepley, Denise M.
;; APPLICANT: Liu, Xiaohong
;; APPLICANT: Mezik, Amanda J.
;; APPLICANT: Padigar, Muralidhara
;; APPLICANT: Patutajan, Meera
;; APPLICANT: Raetelli, Luca
;; APPLICANT: Shen, Lei
;; APPLICANT: Shenoy, Suresh G.
;; APPLICANT: Shinkets, Richard A.
;; APPLICANT: Spaderna, Steven K.
;; APPLICANT: Spytek, Kimberly A.
;; APPLICANT: Szekeres, Edward S. Jr.
;; APPLICANT: Taupier, Raymond J. Jr.
;; APPLICANT: Tchernev, Velizar T.
;; APPLICANT: Zerhusen, Bryan D.
;; APPLICANT: Voss, Edward Z.
;; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;; FILE REFERENCE: 21402-050 CIP
;; CURRENT APPLICATION NUMBER: US/10/190,115
;; CURRENT FILING DATE: 2003-02-10
;; PRIOR APPLICATION NUMBER: 60/303,168
;; PRIOR FILING DATE: 2001-07-05
;; PRIOR APPLICATION NUMBER: 60/366,996
;; PRIOR FILING DATE: 2002-04-01
;; PRIOR APPLICATION NUMBER: 60/386,816
;; PRIOR FILING DATE: 2002-06-07
;; PRIOR APPLICATION NUMBER: 60/215,854
```

```
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585,
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2001-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: CuroSeqList version 0.1
; SEQ ID NO 2
; LENGTH: 2469
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-190-115-2
```

```
Query Match          69.4%; Score 34; DB 15; Length 2469;
Best Local Similarity 57.1%; Pred. No. 1e+03;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 VNDRCT 9
       :|||:|:
Db      1625 IDNRQCV 1631
```

```
RESULT 16
US-10-369-072-2
; Sequence 2, Application US/10369072
; Publication No. US20040014081A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchenev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zehusen, Bryan
; APPLICANT: Batturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szekeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: No. US20040014081A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-17
```

```
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2469
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-072-2
```

```
Query Match          69.4%; Score 34; DB 15; Length 2469;
Best Local Similarity 57.1%; Pred. No. 1e+03;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 VNDRCT 9
       :|||:|:
Db      1625 IDNRQCV 1631
```

```
RESULT 17
US-10-190-115-27
; Sequence 27, Application US/10190115
; Publication No. US20030207394A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Alsobrook, John P. II
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Grosse, William M.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Lepley, Denise M.
; APPLICANT: Liu, Xiaohong
; APPLICANT: Mezick, Amanda J.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Rastelli, Luca
; APPLICANT: Shen, Lei
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Szekeres, Edward S. Jr.
; APPLICANT: Taupier, Raymond J. Jr.
; APPLICANT: Tchenev, Velizar T.
; APPLICANT: Zehusen, Bryan D.
; APPLICANT: Voss, Edward Z.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-050 CIP
; CURRENT APPLICATION NUMBER: US/10/190,115
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/303,168
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585,
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2001-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
```

```
;; PRIOR FILING DATE: 2000-07-17
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 136
;; SOFTWARE: CuraseqList version 0.1
;; SEQ ID NO 27
;; LENGTH: 2471
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-190-115-27

Query Match          69.4%; Score 34; DB 15; Length 2471;
Best Local Similarity 57.1%; Pred. No. 1e+03;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      3 VDNECTI 9
Db      1627 IDNRQCV 1633

RESULT 18
US-10-369-072-27
;; Sequence 27, Application US/10369072
;; Publication No. US20040014081A1
;; GENERAL INFORMATION:
;; APPLICANT: Alsobrook II, John P
;; APPLICANT: Spaderna, Stephen K
;; APPLICANT: Tchenev, Velizar
;; APPLICANT: Liu, Xiaohong
;; APPLICANT: Shenoy, Kimbely
;; APPLICANT: Zehrusen, Bryan
;; APPLICANT: Paturajan, Meera
;; APPLICANT: Taupier, Raymond T
;; APPLICANT: Rastelli, Luca
;; APPLICANT: Grosse, William M
;; APPLICANT: Szerkes, Edward S
;; APPLICANT: Lepley, Denise M
;; APPLICANT: Shen, Lei
;; APPLICANT: Burgess, Catherine E
;; APPLICANT: Shinkets, Richard
;; APPLICANT: Padigaru, Muralidhara
;; TITLE OF INVENTION: No. US20040014081A1e1 Proteins and Nucleic Acids Encoding Same
;; FILE REFERENCE: 21402-050 CON2
;; CURRENT APPLICATION NUMBER: US/10/369,072
;; PRIOR FILING DATE: 2003-02-18
;; PRIOR APPLICATION NUMBER: 10/174,372
;; PRIOR FILING DATE: 2002-06-17
;; PRIOR APPLICATION NUMBER: 09/898,994
;; PRIOR FILING DATE: 2001-07-03
;; PRIOR APPLICATION NUMBER: 60/215,854
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/215,856
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/215,902
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/216,585
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/216,586
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/216,722
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/218,622
;; PRIOR FILING DATE: 2000-07-17
;; PRIOR APPLICATION NUMBER: 60/218,992
;; PRIOR FILING DATE: 2000-07-17
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 100
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 27
;; LENGTH: 2471
;; TYPE: PRT
;; ORGANISM: Drosophila melanogaster
US-10-369-072-27
```

```
Query Match          69.4%; Score 34; DB 15; Length 2471;
Best Local Similarity 57.1%; Pred. No. 1e+03;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      3 VDNECTI 9
Db      1627 IDNRQCV 1633

RESULT 19
US-10-190-115-29
;; Sequence 29, Application US/10190115
;; Publication No. US20030207394A1
;; GENERAL INFORMATION:
;; APPLICANT: Alsobrook, John P. II
;; APPLICANT: Boldog, Ferenc L.
;; APPLICANT: Burgess, Catherine E.
;; APPLICANT: Casman, Stacie J.
;; APPLICANT: Grosse, William M.
;; APPLICANT: Gusev, Vladimir Y.
;; APPLICANT: Ji, Weizhen
;; APPLICANT: Lepley, Denise M.
;; APPLICANT: Liu, Xiaohong
;; APPLICANT: Merick, Amanda J.
;; APPLICANT: Padigaru, Muralidhara
;; APPLICANT: Paturajan, Meera
;; APPLICANT: Rastelli, Luca
;; APPLICANT: Shen, Lei
;; APPLICANT: Shenoy, Suresh G.
;; APPLICANT: Shinkets, Richard A.
;; APPLICANT: Spaderna, Steven K.
;; APPLICANT: Spytek, Kimberly A.
;; APPLICANT: Szerkes, Edward S. Jr.
;; APPLICANT: Taupier, Raymond J. Jr.
;; APPLICANT: Tchenev, Velizar T.
;; APPLICANT: Zehrusen, Bryan D.
;; APPLICANT: Vos8, Edward Z.
;; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;; FILE REFERENCE: 21402-050 CIP
;; CURRENT APPLICATION NUMBER: US/10/190,115
;; CURRENT FILING DATE: 2003-02-10
;; PRIOR APPLICATION NUMBER: 60/303,168
;; PRIOR FILING DATE: 2001-07-05
;; PRIOR APPLICATION NUMBER: 60/368,996
;; PRIOR FILING DATE: 2002-04-01
;; PRIOR APPLICATION NUMBER: 60/386,816
;; PRIOR FILING DATE: 2002-06-07
;; PRIOR APPLICATION NUMBER: 60/215,854
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/215,856
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/215,902
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/216,585,
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/216,586
;; PRIOR FILING DATE: 2001-07-07
;; PRIOR APPLICATION NUMBER: 60/216,722
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/218,622
;; PRIOR FILING DATE: 2000-07-17
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 136
;; SOFTWARE: CuraseqList version 0.1
;; SEQ ID NO 29
;; LENGTH: 2531
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-190-115-29

Query Match          69.4%; Score 34; DB 15; Length 2531;
Best Local Similarity 57.1%; Pred. No. 1e+03;
```

Matches: 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNRECT 9

Db 1670 IDNRQCV 1676

RESULT 20

US-10-369-072-29

Sequence 29, Application US/10369072

Publication No. US20040014081A1

GENERAL INFORMATION:

APPLICANT: Alsbrook II, John P

APPLICANT: Spaderna, Stephen K

APPLICANT: Tchervet, Velizar

APPLICANT: Liu, Xiaohong

APPLICANT: Shenoy, Suresh

APPLICANT: Spytek, Kimberly

APPLICANT: Zehusen, Bryan

APPLICANT: Patnirajan, Meera

APPLICANT: Taupier, Raymond T

APPLICANT: Rastelli, Luca

APPLICANT: Grosse, William M

APPLICANT: Szekeres, Edward S

APPLICANT: Lepley, Denise M

APPLICANT: Shen, Lei

APPLICANT: Burgess, Catherine E

APPLICANT: Shimkets, Richard

APPLICANT: Padigaru, Muraidhara

TITLE OF INVENTION: No. US20040014081A1 Proteins and Nucleic Acids Encoding Same

FILE REFERENCE: 21402-050 CON2

CURRENT APPLICATION NUMBER: US/10/369,072

CURRENT FILING DATE: 2003-02-18

PRIOR FILING DATE: 2002-06-17

PRIOR APPLICATION NUMBER: 09/898,994

PRIOR FILING DATE: 2001-07-03

PRIOR APPLICATION NUMBER: 60/215,854

PRIOR FILING DATE: 2000-07-03

PRIOR APPLICATION NUMBER: 60/215,856

PRIOR FILING DATE: 2000-07-03

PRIOR APPLICATION NUMBER: 60/215,902

PRIOR FILING DATE: 2000-07-03

PRIOR APPLICATION NUMBER: 60/216,585

PRIOR FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: 60/216,586

PRIOR FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: 60/216,722

PRIOR FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: 60/218,622

PRIOR FILING DATE: 2000-07-17

PRIOR APPLICATION NUMBER: 60/218,992

PRIOR FILING DATE: 2000-07-17

PRIOR APPLICATION NUMBER: 60/218,992

PRIOR FILING DATE: 2000-07-17

PRIOR APPLICATION NUMBER: 60/218,992

PRIOR FILING DATE: 2000-07-17

PRIOR APPLICATION NUMBER: 60/218,992

PRIOR FILING DATE: 2000-07-17

PRIOR APPLICATION NUMBER: 60/218,992

PRIOR FILING DATE: 2000-07-17

PRIOR APPLICATION NUMBER: 60/218,992

PRIOR FILING DATE: 2000-07-17

PRIOR APPLICATION NUMBER: 60/218,992

Sequence 12, Application US/10294006

Publication No. US20040013657A1

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Zhou, Ping

APPLICANT: Goodrich, Ryle

APPLICANT: Asundi, Vinod

APPLICANT: Yang, Yonghong

APPLICANT: Zhang, Jie

APPLICANT: Weinman, Tom

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. US20040013657A1 Nucleic Acids and

FILE REFERENCE: 21272-029CIP2H-PCT

CURRENT APPLICATION NUMBER: US/10/294,006

CURRENT FILING DATE: 2002-11-12

PRIOR FILING DATE: 2001-03-22

PRIOR APPLICATION NUMBER: 09/815,925

NUMBER OF SEQ ID NOS: 44

SOFTWARE: pt_FL_genes Version 2.0

SEQ ID NO 12

LENGTH: 2556

TYPE: PRT

ORGANISM: Homo sapiens

US-10-294-006-12

Query Match

Best Local Similarity 57.1%; Score 34; DB 15; Length 2556;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDNRECT 9

Db 1681 IDNRQCV 1687

RESULT 22

US-10-369-493-13322

Sequence 13322, Application US/10369493

Publication No. US2003023675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 13322

LENGTH: 316

TYPE: PRT

ORGANISM: Aspergillus nidulans

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(316)

OTHER INFORMATION: unsure at all Xaa locations

US-10-369-493-13322

Query Match

Best Local Similarity 67.3%; Score 33; DB 15; Length 316;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRE 7

Db 61 AGVDNRQ 67

RESULT 23

RESULT 21

US-10-294-006-12

```
US-10-156-761-14060
; Sequence 14060, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14060
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14060

Query Match      67.3%; Score 33; DB 14; Length 429;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 AGVDNRECI 9
DB      101 SGGNNECV 109

RESULT 24
US-10-354-358-102
; Sequence 102, Application US/10354358
; Publication No. US20030157082A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Hunter, John Joseph
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Teal, Fong-ying
; APPLICANT: Lesoon, Andrea
; APPLICANT: Lightcap, Eric S.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
; TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
; FILE REFERENCE: MP102-020P1R0NM1M
; CURRENT APPLICATION NUMBER: US/10/354,358
; PRIOR FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US 60/353,600
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/364,517
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/371,075
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/371,507
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/372,984
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/374,194
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/382,995
; PRIOR FILING DATE: 2002-05-24
```

```
; PRIOR APPLICATION NUMBER: US 60/385,023
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/388,853
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/389,395
; PRIOR FILING DATE: 2002-06-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 102
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-354-358-102

Query Match      67.3%; Score 33; DB 14; Length 984;
Best Local Similarity 55.6%; Pred. No. 6.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 AGVDNRECI 9
DB      945 AGDTMECV 953

RESULT 25
US-10-116-275-147
; Sequence 147, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: Patent version 3.1
; SEQ ID NO 147
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-147

Query Match      67.3%; Score 33; DB 15; Length 984;
Best Local Similarity 55.6%; Pred. No. 6.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 AGVDNRECI 9
DB      945 AGDTMECV 953

RESULT 26
US-09-901-436A-8
; Sequence 8, Application US/09901436A
; Patent No. US20020098515A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: Cyclostatin I
; FILE REFERENCE: PFI75D4
; CURRENT APPLICATION NUMBER: US/09/901,436A
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/361,737
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 09/023,073
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 08/470,298
; PRIOR FILING DATE: 1995-06-06
```

PRIOR APPLICATION NUMBER: 08/409,731
PRIOR FILING DATE: 1995-03-24
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8
LENGTH: 135
TYPE: PRT
ORGANISM: human
US-09-901-436A-8

Query Match 65.3%; Score 32; DB 9; Length 135;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVDNRECT 9
DB 77 GIDDRKCM 84

RESULT 27
US-10-032-366-3
; Sequence 3, Application US/10032366
; Publication No. US20020173002A1
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Guegler, Karl J.
Shah, Purvu
TITLE OF INVENTION: HUMAN RETINOID BINDING PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/032,366
FILING DATE: 21-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/899,031
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0349 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 50548
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-032-366-3
Query Match 65.3%; Score 32; DB 13; Length 135;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVDNRECT 9

DB 77 GIDDRKCM 84

RESULT 28
US-10-032-366-4
; Sequence 4, Application US/10032366
; Publication No. US20020173002A1
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Guegler, Karl J.
Shah, Purvu
TITLE OF INVENTION: HUMAN RETINOID BINDING PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/032,366
FILING DATE: 21-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/899,031
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0349 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 190948
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-032-366-4
Query Match 65.3%; Score 32; DB 13; Length 135;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVDNRECT 9

DB 77 GIDDRKCM 84

RESULT 29
US-10-171-311-200
; Sequence 200, Application US/10171311
; Publication No. US20030087270A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
Chen, Yan
APPLICANT: Zhao, Xumei
APPLICANT: Monahan, John
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Glatt, Karen

```
;; APPLICANT: Gannavarapu, Manjula
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
;; FILE REFERENCE: MRI-035
;; CURRENT APPLICATION NUMBER: US/10/171,311
;; PRIOR FILING DATE: 2002-06-12
;; PRIOR APPLICATION NUMBER: US 60/298,159
;; PRIOR FILING DATE: 2001-06-13
;; PRIOR APPLICATION NUMBER: US 60/298,155
;; PRIOR FILING DATE: 2001-06-13
;; PRIOR APPLICATION NUMBER: US 60/335,936
;; PRIOR FILING DATE: 2001-11-14
;; NUMBER OF SEQ ID NOS: 238
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 200
;; LENGTH: 135
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-171-311-200

Query Match      65.3%; Score 32; DB 14; Length 135;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      2 GVDNRECT 9
Db      77 GIDDRKCM 84

RESULT 30
US-10-264-049-3046
;; Sequence 3046, Application US/10264049
;; Publication No. US20040005579A1
;; GENERAL INFORMATION:
;; APPLICANT: Birse et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PA13P1
;; CURRENT APPLICATION NUMBER: US/10/264,049
;; PRIOR FILING DATE: 2002-10-04
;; PRIOR APPLICATION NUMBER: PCT/US01/18569
;; PRIOR FILING DATE: 2001-06-07
;; PRIOR APPLICATION NUMBER: US 60/209,467
;; PRIOR FILING DATE: 2000-06-07
;; NUMBER OF SEQ ID NOS: 4360
;; SOFTWARE: PatentIn Ver. 3.1
;; SEQ ID NO 3046
;; LENGTH: 219
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-264-049-3046

Query Match      65.3%; Score 32; DB 15; Length 219;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      2 GVDNRECT 9
Db      161 GIDDRKCM 168

RESULT 31
US-10-369-493-12041
;; Sequence 12041, Application US/10369493
;; Publication No. US20030233675A1
;; GENERAL INFORMATION:
;; APPLICANT: Cao, Yongwei
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Chen, Xianfeng
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
```

```
;; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
;; FILE REFERENCE: 38-10(52052)B
;; CURRENT APPLICATION NUMBER: US/10/369,493
;; PRIOR FILING DATE: 2003-02-28
;; PRIOR APPLICATION NUMBER: US 60/360,039
;; PRIOR FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 12041
;; LENGTH: 221
;; TYPE: PRT
;; ORGANISM: Mesorhizobium loti
US-10-369-493-12041

Query Match      65.3%; Score 32; DB 15; Length 221;
Best Local Similarity 62.5%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 GVDNRECT 9
Db      152 GVDARCT 159

RESULT 32
US-10-369-493-2249
;; Sequence 2249, Application US/10369493
;; Publication No. US20030233675A1
;; GENERAL INFORMATION:
;; APPLICANT: Cao, Yongwei
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Chen, Xianfeng
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; FILE REFERENCE: 38-10(52052)B
;; CURRENT APPLICATION NUMBER: US/10/369,493
;; PRIOR FILING DATE: 2003-02-28
;; PRIOR APPLICATION NUMBER: US 60/360,039
;; PRIOR FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 2249
;; LENGTH: 308
;; TYPE: PRT
;; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2249

Query Match      65.3%; Score 32; DB 15; Length 308;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 VDNRECT 9
Db      24 VDNSECV 30

RESULT 33
US-09-989-545-17
;; Sequence 17, Application US/09989545
;; Patent No. US20020164697A1
;; GENERAL INFORMATION:
;; APPLICANT: Lehner, Sophie
;; APPLICANT: Manning, Stephen
;; APPLICANT: Coyle, Anthony J.
;; APPLICANT: Gutierrez-Ramos, Jose-Carlos
;; TITLE OF INVENTION: No. US20020164697A1 Th2-Specific Molecules and Uses Thereof
;; FILE REFERENCE: 5800-10B
;; CURRENT APPLICATION NUMBER: US/09/989,545
;; PRIOR FILING DATE: 2001-11-20
;; PRIOR APPLICATION NUMBER: 09/168,229
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 09/258,670
;; PRIOR FILING DATE: 1999-02-26
;; NUMBER OF SEQ ID NOS: 24
```

SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Trypanosoma brucei
US-09-989-545-17

Query Match 65.3%; Score 32; DB 9; Length 442;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRE 7
|||:||||
DB 209 AGVDNRE 215

RESULT 34
US-10-446-520-14
; Sequence 14, Application US/10446520
; Publication No. US20030235898A1
; GENERAL INFORMATION:
; APPLICANT: Kloeck, Andrew P.
; APPLICANT: Williams, Deryck J.
; APPLICANT: Salmon, Brandy
; APPLICANT: McAlaird, Merry B.
; TITLE OF INVENTION: NEMATODE GS-LIKE SEQUENCES
; FILE REFERENCE: 12557-020001
; CURRENT APPLICATION NUMBER: US/10/446,520
; PRIOR FILING DATE: 2003-05-27
; PRIOR APPLICATION NUMBER: US 10/098,602
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/276,621
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-446-520-14

Query Match 65.3%; Score 32; DB 15; Length 457;
Best Local Similarity 71.4%; Pred. No. 4.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVDNREC 8
|:|||||
DB 333 GVDNREC 339

RESULT 35
US-10-369-493-9643
; Sequence 9643, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9643
; LENGTH: 787
; TYPE: PRT
; ORGANISM: Desulfitobacterium hafnense
US-10-369-493-9643

Query Match 65.3%; Score 32; DB 15; Length 787;
Best Local Similarity 75.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVDNRECI 9
|||:||||
DB 754 GVDNRECI 761

RESULT 36
US-09-965-536A-10
; Sequence 10, Application US/09965536A
; Publication No. US20030027323A1
; GENERAL INFORMATION:
; APPLICANT: FEDER, J. N.
; APPLICANT: MINTIER, G.
; APPLICANT: RAMANATHAN, C. S.
; APPLICANT: HAWKEN, D. R.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMV5,
; TITLE OF INVENTION: EXPRESSED HIGHLY IN BRAIN AND OVARIAN TISSUES
; FILE REFERENCE: D0041NP
; CURRENT APPLICATION NUMBER: US/09/965,536A
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,713
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/261,781
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/306,605
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/310,436
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Lymnaea stagnalis
US-09-965-536A-10

Query Match 65.3%; Score 32; DB 10; Length 1115;
Best Local Similarity 71.4%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVDNREC 8
|:|||||
DB 107 GVDNREC 113

RESULT 37
US-10-222-668-3
; Sequence 3, Application US/10222668
; Publication No. US2003008884A1
; GENERAL INFORMATION:
; APPLICANT: Heu, Sheau Yu
; APPLICANT: Hsueh, Aaron
; TITLE OF INVENTION: Mammalian Relaxin Receptor
; FILE REFERENCE: STAN-239 WO
; CURRENT APPLICATION NUMBER: US/10/222,668
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/313,259
; PRIOR FILING DATE: 2002-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: H. sapiens
US-10-222-668-3

Query Match 65.3%; Score 32; DB 14; Length 1115;
Best Local Similarity 71.4%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 2 GVNREC 8
Db 107 GLDEREC 113

RESULT 38

US-10-205-823-273
; Sequence 273, Application US/10205823
; Publication No. US20030108963a1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarepu, Manjula
; APPLICANT: Gorbatcheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatc, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 273
; LENGTH: 1285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-273

Query Match 65.3%; Score 32; DB 14; Length 1285;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Cy 3 VDNRECI 9
Db 509 VDNODCI 515

RESULT 39
US-10-177-293-317
; Sequence 317, Application US/10177293
; Publication No. US20030124128a1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatc, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarepu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Baer Jr., Robert C.

APPLICANT: Hortobagyi, Gabriel N.
APPLICANT: Puetzai, Lajos
APPLICANT: Meric, Funda
APPLICANT: Sahin, Aysegul
APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 317
; LENGTH: 1285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-317

Query Match 65.3%; Score 32; DB 14; Length 1285;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Cy 3 VDNRECI 9
Db 509 VDNODCI 515

RESULT 40
US-09-864-761-48480
; Sequence 48480, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48480
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011363.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.99
; OTHER INFORMATION: EST HUMAN HIT: AW024552.1, EVALU8.00e-25
; OTHER INFORMATION: SWISSPROT HIT: OS4827, EVALU8.2.00e-29
US-09-864-761-48480

Query Match 63.3%; Score 31; DB 9; Length 92;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VDNRECT 9
: : : : :
Db 83 INNRECT 89

Search completed: March 3, 2004, 10:27:17
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 09:46:53 ; Search time 50.5 Seconds
(without alignments)
50.355 Million cell updates/sec

Title: US-09-980-064-2
Perfect score: 49
Sequence: 1 AGVDNRECT 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1990s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	9	AA45906	Ab45906 HPV16 L1
2	49	100.0	9	AA46247	Ab46247 Influenza
3	49	100.0	9	AA46260	Ab46260 HPV type
4	49	100.0	20	ABP47241	Abp47241 HPV16 L1
5	49	100.0	497	AA71464	Aa71464 Chimeric
6	49	100.0	499	AA71465	Aa71465 Chimeric
7	49	100.0	505	AA38808	Aa38808 HPV1 L1
8	49	100.0	505	AA38807	Aa38807 BPV1 L1
9	49	100.0	505	AA39903	Aa39903 Human pap
10	49	100.0	505	AA63116	Aa63116 Human pap
11	49	100.0	505	AA53487	Aa53487 Human pap
12	49	100.0	505	AA53486	Aa53486 Human pap
13	49	100.0	505	AA44142	Aa44142 Human pap
14	49	100.0	505	AA44143	Aa44143 Human pap
15	49	100.0	505	AA54034	Aa54034 HPV16 L1
16	49	100.0	505	AA54073	Aa54073 HPV16 L1
17	49	100.0	505	AA96747	Aa96747 Human pap
18	49	100.0	505	AA96748	Aa96748 Human pap
19	49	100.0	505	AA08019	Aa08019 Human pap
20	49	100.0	505	AA53580	Aa53580 HPV16 L1
21	49	100.0	505	AA53579	Aa53579 HPV16 L1
22	49	100.0	505	AA57720	Aa57720 Human pap
23	49	100.0	505	AB77479	Ab77479 HPV16-L1
24	49	100.0	505	AB77478	Ab77478 HPV16-L1
25	49	100.0	505	AB75800	Ab75800 Human pap

26	49	100.0	505	6	ABG75799	Abg75799 Human pap
27	49	100.0	505	6	ADA27363	Ada27363 HPV-16 L1
28	49	100.0	505	7	ADA92541	Ada92541 HPV-16 L1
29	49	100.0	505	7	ADA14290	Ada14290 HPV-16 L1
30	49	100.0	505	7	AAE38615	Aae38615 HPV-16 L1
31	49	100.0	505	7	ADC26174	Adc26174 HPV16 L1
32	49	100.0	505	7	ADC26175	Adc26175 HPV16 L1
33	49	100.0	531	3	AA823925	Aa823925 Human pap
34	49	100.0	531	4	AA898422	Aa898422 Human pap
35	49	100.0	533	5	ABB77482	Abb77482 HPV16-L1/
36	49	100.0	533	5	ABB77481	Abb77481 HPV16-L1/
37	49	100.0	542	5	ABB77483	Abb77483 HPV16-L1/
38	41	83.7	504	4	AA898436	Aa898436 Human pap
39	39	79.6	20	2	AA804969	Aa804969 Papilloma
40	39	79.6	51	2	AA832059	Aa832059 Human Igg
41	39	79.6	51	2	AA832062	Aa832062 Human Iga
42	39	79.6	499	4	AA898449	Aa898449 Human pap
43	37	75.5	110	6	ADB11572	Adb11572 Al10iococ
44	37	75.5	185	4	AAU47795	Aau47795 Propionib
45	37	75.5	185	6	ABM44314	Abm44314 Propionib

ALIGNMENTS

RESULT 1
AA45906 standard; peptide; 9 AA.

AA45906; 20-MAR-2001 (first entry)

HPV16 L1 protein P-12 peptide fragment.

In-vitro detection; antigen-specific immune response; vaccine; capsomer; capsid; virus-like particle; antigen-presenting target cell; chimeric papillomavirus-like particle; autoimmune disease.

Human papillomavirus.

WO200073790-A1.

07-DEC-2000.

31-MAY-2000; 2000WO-EP005003.

01-JUN-1999; 99DB-01025234.

(MEDI-) MEDIGENE AG.

Jochmus I, Nleland J;

WPI; 2001-049960/06.

Detecting antigen-specific immune response, comprising capsomer, capsid, VLP (virus like particles) and/or CVP (chimeric papillomavirus-like particles), pre-incubated antigen-presenting target cell and effector cell, useful in diagnosis.

Example 1; Page 31; 71pp; German.

This invention describes a novel test system for in vitro detection of an antigen-specific immune response comprising at least 1 component (1) selected from capsomers, stable capsomers, capsids, virus-like particles (VLP's) and/or chimeric papillomavirus-like particles (CVP's) and at least 1 antigen-presenting target cell which has been incubated with (1) and effector cells from the immune system. The test systems are useful in the in vitro detection of activation of effector cells of the immune system by (1), specifically for: (1) quality control of prophylactic and/or therapeutic vaccines containing (1) or cells; (2) identification of epitopes, peptides or protein fragments which elicit an immune response, especially a cellular immune response; (3) monitoring the

CC immune status of an organism towards a stimulus (especially a stimulus
CC which is difficult to detect); (4) monitoring a vaccination; (5)
CC identifying HLA (human leukocyte antigen)-haplotypes which mediate
CC immunity towards a specific stimulus; (6) differentiating and
CC characterizing autoimmune diseases with respect to different autoimmune
CC antigens; or (7) differentiating tumor types with respect to different
CC tumor antigens. The test system is easier and less expensive to handle
CC than prior art systems based on vaccinia and adenoviruses, and allows
CC quantitative determination of the immune response of cytotoxic T-cells.
CC Differentiation between immune cells is possible. The uptake of antigens
CC in the cells is independent of major histocompatibility complex
CC molecules. No viral infection associated with viral protein expression
CC and viral replication takes place. The system can be standardized
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 49; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
| | | | |
| | | | |
DB 1 AGVDNRECI 9

RESULT 2
AAB46247
ID AAB46247 standard; peptide; 9 AA.

XX AAB46247;

DT 04-APR-2001 (first entry)

XX Influenza virus cytotoxic T-cell epitope SEQ ID NO 2.

XX Cytotoxic T cell; epitope; LI protein; antiviral; antitumor; antigen;

KW vaccine; tumor; protective immune response.

XX Influenza virus.

XX DE19925235-A1.

XX 07-DEC-2000.

XX 01-JUN-1999; 99DE-01025235.

XX 01-JUN-1999; 99DE-01025235.

PA (MEDI-) MEDIGENE AG.

PA (DKFZ-) DKFZ DEUT KREBSFORSCHUNGSZENTRUM.

PI Schaefer K, Faath S, Jochmus I, Nieldand J, Osen W;

XX WPI; 2001-062580/08.

PT New T cell epitopes from the papilloma virus LI protein, useful for
PT detecting or inducing an immune response, e.g. in vaccines.

PS Claim 1; Page 12; 26pp; German.

CC This invention describes novel T cell epitopes (A) of the human papilloma
CC virus LI protein. The invention also describes (a) compounds (I)
CC containing an (A) provided it is not a natural LI protein nor an
CC exclusively N- or C-terminal deletion variant of natural LI; (b) nucleic
CC acid (II) that encodes (A) or (I); (c) vector, particularly an expression
CC vector, that contains (II); (d) cells that contain, and preferably
CC present, (A); (e) complex of (A) or (I) with at least one additional
CC molecule (III); in vitro detection of T cell activation by a compound
CC containing at least one (A); (f) preparing cells of (d); and (g) test
CC system for in vitro detection of T cells. The products of the invention
CC have antiviral and antitumor activity. (A), also compounds or complexes
CC containing it or nucleic acid or vectors encoding it, is used for
CC detection of an immune response, particularly detecting LI-specific

CC cytotoxic T cells or to determine the LI protein-specific antigenicity of
CC compounds and complexes that contain (A), and for inducing a protective
CC immune response against papilloma virus and related tumors, particularly
CC as vaccines

SQ Sequence 9 AA;

Query Match 100.0%; Score 49; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
| | | | |
| | | | |
DB 1 AGVDNRECI 9

RESULT 3
AAB46260
ID AAB46260 standard; peptide; 9 AA.

XX AAB46260;

DT 04-APR-2001 (first entry)

XX HPV type 16 cytotoxic T-cell epitope SEQ ID NO 15.

XX Cytotoxic T cell; epitope; LI protein; antiviral; antitumor; antigen;

KW vaccine; tumor; protective immune response.

XX Human Papillomavirus.

XX DE19925235-A1.

XX 07-DEC-2000.

XX 01-JUN-1999; 99DE-01025235.

XX 01-JUN-1999; 99DE-01025235.

PA (MEDI-) MEDIGENE AG.

PA (DKFZ-) DKFZ DEUT KREBSFORSCHUNGSZENTRUM.

PI Schaefer K, Faath S, Jochmus I, Nieldand J, Osen W;

XX WPI; 2001-062580/08.

PT New T cell epitopes from the papilloma virus LI protein, useful for
PT detecting or inducing an immune response, e.g. in vaccines.

PS Example 4; Page 17; 26pp; German.

CC This invention describes novel T cell epitopes (A) of the human papilloma
CC virus LI protein. The invention also describes (a) compounds (I)
CC containing an (A) provided it is not a natural LI protein nor an
CC exclusively N- or C-terminal deletion variant of natural LI; (b) nucleic
CC acid (II) that encodes (A) or (I); (c) vector, particularly an expression
CC vector, that contains (II); (d) cells that contain, and preferably
CC present, (A); (e) complex of (A) or (I) with at least one additional
CC molecule (III); in vitro detection of T cell activation by a compound
CC containing at least one (A); (f) preparing cells of (d); and (g) test
CC system for in vitro detection of T cells. The products of the invention
CC have antiviral and antitumor activity. (A), also compounds or complexes
CC containing it or nucleic acid or vectors encoding it, is used for
CC detection of an immune response, particularly detecting LI-specific
CC cytotoxic T cells or to determine the LI protein-specific antigenicity of
CC compounds and complexes that contain (A), and for inducing a protective
CC immune response against papilloma virus and related tumors, particularly
CC as vaccines

SQ Sequence 9 AA;

Query Match 100.0%; Score 49; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;

```

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGVDNRECT 9
Db 1 AGVDNRECT 9

RESULT 4
ABP47241
ID ABP47241 standard; peptide; 20 AA.
XX
AC ABP47241;
XX
DT 29-AUG-2003 (revised)
DT 07-AUG-2003 (revised)
DT 15-AUG-2002 (first entry)
XX
DE HPV16 L1 T-cell epitope peptide P13 SEQ ID NO 13.
XX
KM HPV; human papillomavirus; T-cell epitope; L1; E7; virucide; cytostatic;
KM vaccine; immune response.
XX
OS Human papillomavirus type 16.
XX
PN WO200244384-A2.
XX
PD 06-JUN-2002.
XX
PF 30-NOV-2001; 2001WO-EP014037.
XX
PR 01-DEC-2000; 2000DE-01059631.
XX
PA (MEDI-) MEDIGENE AG.
PI Nieland J, Kaufmann A;
XX
XX WPI; 2002-471833/50.
XX
DR New T cell epitopes derived from human papilloma virus, useful for
PT detecting immune activation and in a vaccine against papilloma virus.
XX
PS Example 1; Page 35; 126pp; German.
XX
CC The invention relates to T-cell epitopes (I) derived from human
CC papillomavirus (HPV). (I) have virucide and cytostatic activity and are
CC able to induce a cytotoxic T cell response or mediate T helper cell
CC function. (II) and also compounds (II) containing them, vectors that
CC express (I) or (II), cells that contain (I) or (II) and/or complexes of
CC (I) or (II) with another compound, are used to detect or induce an immune
CC response, for diagnosis or therapy, particularly in vaccines for control
CC of human papilloma virus infection. The present sequence is that of a HPV
CC T-cell epitope of the invention. (Updated on 07-AUG-2003 to correct OS
CC field.) (Updated on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 20 AA;
Qy Query Match 100.0%; Score 49; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 AGVDNRECT 9
7 AGVDNRECT 15

RESULT 5
AA71464
ID AA71464 standard; protein; 497 AA.
XX
AC AA71464;
XX
DT 04-OCT-2000 (first entry)
XX

```

```

DE Chimeric biotin-binding human papillomavirus mutant L1 protein-168.
XX
KM Human papillomavirus; HPV, L1 protein; L2 protein; biotin; chimeric;
KM cytostatic; antiviral; gene therapy; vaccine; capsomere; VLP;
KM virus-like particle; viral antigen; tumour; cytotoxin; cytokine; IL;
KM interleukin; chemotherapeutic agent; radioactive agent; mutant; mutein.
XX
OS Human papillomavirus type 16.
OS unidentified.
OS Chimeric.
XX
XX Key Location/Qualifiers
XX Region 1..483
XX Binding-site /label= HPV_16_mutant_L1_protein
XX FT 484..497
XX FT /label= Biotin_binding_peptide
XX FT 484..486
XX FT /note= "(G1y)3 hinge"
XX
XX WO200031128-A1.
XX
XX 02-JUN-2000.
XX
XX 22-NOV-1999; 99WO-US027555.
XX
XX 23-NOV-1998; 98US-0109510P.
XX 06-OCT-1999; 99US-00413611.
XX
XX (LOYO) UNIV LOYOLA CHICAGO.
XX
PI Mueller M, Kast WM, Nieland JD, Velders MP;
XX
XX WPI; 2000-400041/34.
XX
DR N-PSDB; AAD01237.
XX
XX Chimeric protein comprising a papillomavirus L1 or L2 protein and a
XX biotin-binding polypeptide, useful for delivering substances such as
XX proteins, nucleic acids and lipids into cells, particularly
XX papillomavirus infected cells.
XX
XX Example 1; Page 18-20; 27pp; English.
XX
XX The patent discloses a chimeric protein comprising papillomavirus L1 or
XX L2 protein and a biotin-binding polypeptide. Capsomere, papillomavirus or
XX virus-like particle (VLP) comprising the chimeric protein is used for
XX delivering a wide variety of biotinylated compounds e.g. proteins,
XX nucleic acids and lipids into cells, particularly papillomavirus infected
XX cells. The chimeric protein may also be used as a vaccine when the
XX biotinylated substance is a viral antigen e.g. papillomavirus E2 or E7
XX proteins. It may be useful for treating tumours or other papillomavirus-
XX related lesions when the substance is a cytotoxin, chemotherapeutic
XX agent, radioactive agent, or a gene encoding a cytokine or interleukin.
XX The present sequence is a chimeric protein 168 consisting of human
XX papillomavirus 16 (HPV-16) mutant L1 protein fused to a biotin-binding
XX peptide at the C-terminus. The L1 protein was derived from the wild-type
XX sequence by deleting the amino acid Cys428. The modification was done to
XX prevent assembly of VLPs while allowing the production of capsomeres at
XX high efficiencies
SQ Sequence 497 AA;
Qy Query Match 100.0%; Score 49; DB 3; Length 497;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 AGVDNRECT 9
139 AGVDNRECT 147

RESULT 6
AA71465
ID AA71465 standard; protein; 499 AA.

```

```

XX AC AAY71465;
XX DT 04-OCT-2000 (first entry)
XX DE Chimeric biotin-binding human papillomavirus mutant L1 protein-169.
XX KW Human papillomavirus; HPV; L1 protein; L2 protein; biotin; chimeric;
XX KW cytostatic; antiviral; gene therapy; vaccine; capsomere; VLP;
XX KW virus-like particle; viral antigen; tumour; cytotoxin; cytokine; IL;
XX KW interleukin; chemotherapeutic agent; radioactive agent; mutant; mteuin.
XX OS Human papillomavirus type 16.
XX OS Unidentified.
XX OS Chimeric.
XX FT Key Location/Qualifiers
XX FT Region 1..483
XX FT Binding-site /label= HPV_16_mutant_L1_protein
XX FT /label= Biotin_binding_domain
XX FT Region 484..499
XX FT /note= "(Gly)3 hinge"
XX PN WO200031128-A1.
XX PD 02-JUN-2000.
XX PF 22-NOV-1999; 99WO-US027555.
XX PR 23-NOV-1998; 98US-0109510P.
XX PR 06-OCT-1999; 99US-00413611.
XX PA (LOYO ) UNIV LOYOLA CHICAGO.
XX PI Mueller M, Kast WM, Nieland JD, Velders MP;
XX DR WPI; 2000-400041/34.
XX DR N-PSDB; AAD01238.
XX PT Chimeric protein comprising a papillomavirus L1 or L2 protein and a
XX PT biotin-binding polypeptide, useful for delivering substances such as
XX PT proteins, nucleic acids and lipids into cells, particularly
XX PT papillomavirus infected cells.
XX PS Example 1; Page 23-24; 27pp; English.
XX CC The patent discloses a chimeric protein comprising papillomavirus L1 or
XX CC L2 protein and a biotin-binding polypeptide. Capsomere, papillomavirus or
XX CC virus-like particle (VLP) comprising the chimeric protein is used for
XX CC delivering a wide variety of biotinylated compounds e.g. proteins,
XX CC nucleic acids and lipids into cells, particularly papillomavirus infected
XX CC cells. The chimeric protein may also be used as a vaccine when the
XX CC biotinylated substance is a viral antigen e.g. papillomavirus E2 or E7
XX CC proteins. It may be useful for treating tumours or other papillomavirus-
XX CC related lesions when the substance is a cytotoxin, chemotherapeutic
XX CC agent, radioactive agent, or a gene encoding a cytokine or interleukin.
XX CC The present sequence is a chimeric protein 169 consisting of human
XX CC papillomavirus 16 (HPV-16) mutant L1 protein fused to a biotin-binding
XX CC peptide at the C-terminus. The L1 protein was derived from the wild-type
XX CC sequence by deleting the amino acid Cys428. The modification was done to
XX CC prevent assembly of VLPs while allowing the production of capsomeres at
XX CC high efficiencies
XX SQ Sequence 499 AA;

```

```

Query Match 100.0%; Score 49; DB 3; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AGVDNRECI 9
DB 139 AGVDNRECI 147

```

```

RESULT 7
AA38808
ID AAR38808 standard; protein; 505 AA.
XX AC AAR38808;
XX DT 25-MAR-2003 (revised)
XX DT 17-DEC-2001 (revised)
XX DT 21-JAN-1994 (first entry)
XX DE HPV1 L1.
XX KW L1; capsid protein; bovine; human; papillomavirus; baculovirus;
XX KW transfer vector; promoter; capsid protein; transformation; vaccine;
XX KW neutralising antibody; vertebrate.
XX OS Human papillomavirus.
XX PN USN8032869-N.
XX PD 15-JUL-1993.
XX PF 16-MAR-1993; 93US-00032869.
XX PR 03-SEP-1992; 92US-00941371.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX PI Lowy DR, Schiller JT, Kirnbauer R;
XX DR WPI; 1993-249995/31.
XX DR N-PSDB; AAQ47167.
XX PT Recombinant papilloma virus capsid proteins - for vaccines against
XX PT papilloma virus and for diagnosis of virus infection.
XX PS Example 1; Page 35-37; 45pp; English.
XX CC The sequences given in AAR38807-08 represent the L1 capsid proteins from
XX CC bovine and human papillomavirus respectively. The DNA encoding these
XX CC sequences may be inserted into a baculovirus transfer vector and
XX CC operatively expressed by a promoter of the vector, and the capsid protein
XX CC produced by transformed cells. These capsid proteins may be used in
XX CC vaccines to induce high-titre neutralising antibody response in
XX CC vertebrates. (Note: Revised entry submitted to correct the patent number
XX CC format of US Government-owned NTIS applications to prevent clashes with
XX CC ongoing US granted patent numbers. For further information please visit
XX CC the Derwent web site at www.derwent.com/dwpi/updates/ntis_us.html.)
XX CC (Updated on 25-MAR-2003 to correct PF field.)
XX SQ Sequence 505 AA;

```

```

Query Match 100.0%; Score 49; DB 2; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AGVDNRECI 9
DB 139 AGVDNRECI 147

```

```

RESULT 8
AAR38807
ID AAR38807 standard; protein; 505 AA.
XX AC AAR38807;
XX DT 25-MAR-2003 (revised)
XX DT 17-DEC-2001 (revised)
XX DT 21-JAN-1994 (first entry)
XX

```

DE BPV1 L1.
 XX LI; capsid protein; bovine; human; papillomavirus; baculovirus;
 KW transfer vector; promoter; capsid protein; transformation; vaccine;
 KW neutralising antibody; vertebrate.
 XX
 OS Bovine papillomavirus.
 XX
 FN USN8032869-N.
 XX
 PD 15-JUL-1993.
 XX
 PF 16-MAR-1993; 93US-00032869.
 XX
 PR 03-SEP-1992; 92US-00941371.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 XX
 PI Lowy DR, Schiller JT, Kirnbauer R;
 XX
 DR WPI; 1993-249995/31.
 DR N-PSDB; AAQ47166.
 XX
 PT Recombinant papilloma virus capsid proteins - for vaccines against
 PT papilloma virus and for diagnosis of virus infection.
 XX
 PS Example 1; Page 32-34; 45pp; English.
 XX
 CC The sequences given in AAR38807-08 represent the L1 capsid proteins from
 CC bovine and human papillomavirus respectively. The DNA encoding these
 CC sequences may be inserted into a baculovirus transfer vector and
 CC operatively expressed by a promoter of the vector, and the capsid protein
 CC produced by transformed cells. These capsid proteins may be used in
 CC vaccines to induce high-titre neutralising antibody response in
 CC vertebrates. (Note: Revised entry submitted to correct the patent number
 CC format of US Government-owned NTIS applications to prevent clashes with
 CC ongoing US granted patent numbers. For further information please visit
 CC the Derwent web site at www.derwent.com/dwpl/updates/ntis_us.html.)
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC
 SQ Sequence 505 AA;
 Query Match 100.0%; Score 49; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;
 Oy 1 AGVDNRECT 9
 Db 139 AGVDNRECT 147
 RESULT 9
 AAW39903
 ID AAW39903 standard; protein; 505 AA.
 XX
 AC AAW39903;
 XX
 DT 26-MAY-1998 (first entry)
 XX
 DE HPV16 L1 capsid protein.
 XX
 KW Capsid protein; L1; HPV16; vaccine; prevention; treatment; self-assembly;
 KW viral protein; capsomer; capsid; antigenic epitope.
 XX
 OS Human papillomavirus.
 XX
 FN Key Location/Qualifiers
 FT 1..505
 FT /label= L1
 FT /note= "Partial capsid protein"
 XX
 PN US5716620-A.
 XX

PD 10-FEB-1998.
 XX
 PF 07-JUN-1995; 95US-00475783.
 XX
 PR 03-SEP-1992; 92US-00941371.
 PR 16-MAR-1993; 93US-00032869.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Kirnbauer R, Lowy DR, Schiller JT;
 XX
 DR WPI; 1998-158363/14.
 DR N-PSDB; AAV03855.
 XX
 PT Vaccine against human papilloma virus - comprises HPV16 L1 polypeptide.
 PT
 PS Disclosure; Col 17-20; 20pp; English.
 XX
 CC This sequence represents the L1 capsid protein from Human Papillomavirus
 CC strain 16 (HPV16). A recombinant form of this viral protein which is
 CC capable of self-assembly into capsomer structures and viral capsids that
 CC comprise conformational antigenic epitopes can be used as a vaccine for
 CC the prevention or treatment of papillomavirus infections in vertebrates.
 CC The vaccine comprises a unit dose of a composition containing a self
 CC assembled HPV16 with at least 1 papillomavirus L1 conformational epitope
 CC
 SQ Sequence 505 AA;
 Query Match 100.0%; Score 49; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;
 Oy 1 AGVDNRECT 9
 Db 139 AGVDNRECT 147
 RESULT 10
 AAW63116
 ID AAW63116 standard; protein; 505 AA.
 XX
 AC AAW63116;
 XX
 DT 08-OCT-1998 (first entry)
 XX
 DE Human papillomavirus type 16 VLP protein D202E mutant.
 XX
 KW Human papillomavirus type 16; HPV16; virus-like particle; VLP;
 KW HPV16D202E; H16.U4 antibody; H16.V5 antibody; identification;
 KW HPV16 type; characterisation; HPV type 16 infection.
 XX
 OS Synthetic.
 OS Human papillomavirus.
 XX
 FN Key Location/Qualifiers
 FT Misc-difference 202
 FT /label= D202E
 FT
 PD WO9825646-A1.
 PD 18-JUN-1998.
 XX
 PF 05-DEC-1997; 97WO-US022023.
 XX
 PR 09-DEC-1996; 96US-0032633P.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Ludmerer S;
 XX
 DR WPI; 1998-348266/30.
 DR
 PT Human papillomavirus type 16 D202E virus-like particles - used for

PT characterizing HPV16 type infections.
 XX
 PS Claim 1; Page; 37pp; English.
 CC
 CC The present sequence represents a synthetic human papillomavirus type 16
 CC (HPV16) virus-like particle (VLP) which contains a Glu substitution for
 CC the naturally occurring Asp at position 202. The protein is designated
 CC HPV16Dd202E. The novel VLP HPV:202E binds H16.V4 antibody but not H16.V5
 CC antibody and thus allows specific identification of HPV-type 16 infections. note:
 CC VLPs are used in the characterisation of HPV-type 16 infections. note:
 CC the sequence does not appear in the specification. It was created using
 CC the sequence given in GenBank accession number AF084952, and information
 CC provided in the specification
 CC
 SQ Sequence 505 AA;
 XX
 Query Match 100.0%; Score 49; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGVDNRECI 9
 DB 139 AGVDNRECI 147
 XX
 RESULT 11
 AAW53487
 ID AAW53487 standard; protein; 505 AA.
 XX
 AC AAW53487;
 XX
 DT 10-JUL-1998 (first entry)
 XX
 DE Human papillomavirus proto-type HPV 16 L1 genome 5637-7155 protein.
 XX
 XX Human papillomavirus; HPV 16; L1 gene; immunisation; capsid;
 KM conformational epitope; vaccine; sequelae; vertebrate.
 XX
 OS Human papillomavirus.
 XX
 PN US5744142-A.
 XX
 PD 28-APR-1998.
 XX
 PF 07-JUN-1995; 95US-00475782.
 XX
 PR 03-SEP-1992; 92US-00941371.
 PR 16-MAR-1993; 93US-00032869.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Schiller JT, Kirnbauer R, Lowy DR;
 XX
 DR WPI; 1998-271010/24.
 DR N-PSDB; AAV25818.
 XX
 PT Immunisation of mammals and humans against papillomavirus infection -
 PT comprises administering recombinant self-assembled capsid proteins
 PT containing conformational epitopes.
 XX
 PS Disclosure; Col 17-20; 20pp; English.
 XX
 CC The present sequence represents the human papillomavirus proto-type HPV
 CC 16 L1 genome 5637-7155 protein, from the present invention. The present
 CC invention describes the immunisation of a host mammal against a
 CC papillomavirus comprising administering to the host, according to an
 CC immunising schedule, a self-assembled Human Papilloma Virus (HPV16)
 CC capsid (I) containing at least 1 papillomavirus L1 conformational
 CC epitope. The papillomavirus L1 conformational epitopes are produced by
 CC letting a genetic construct comprising a papillomavirus L1 gene direct
 CC recombinant expression of the conformational epitope in a transformed
 CC eukaryotic host cell by self-assembly of papillomavirus capsids
 CC containing a L1 polypeptide having an amino acid sequence encoded by the

CC nucleic acid sequence given in the specification (see AAV25817). (I) are
 CC useful for the diagnosis of and as vaccines for the prevention of
 CC papillomavirus infections and their benign and malignant sequelae in
 CC vertebrates. Recombinant conformed papillomavirus proteins are provided
 CC which can be used to produce renewable papillomavirus reagents of any
 CC selected species and type in cell culture. The self-assembled recombinant
 CC L1 capsid protein has the efficacy of intact papillomavirus particles to
 CC induce high levels of neutralising antiserum, in contrast to prior art L1
 CC protein extracted from recombinant bacteria or denatured virions
 CC
 SQ Sequence 505 AA;
 XX
 Query Match 100.0%; Score 49; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGVDNRECI 9
 DB 139 AGVDNRECI 147
 XX
 RESULT 12
 AAW53486
 ID AAW53486 standard; protein; 505 AA.
 XX
 AC AAW53486;
 XX
 DT 10-JUL-1998 (first entry)
 XX
 DE Human papillomavirus wild-type HPV 16 L1 conformational epitope.
 XX
 XX Human papillomavirus; HPV 16; L1 gene; immunisation; capsid;
 KM conformational epitope; vaccine; sequelae; vertebrate.
 XX
 OS Human papillomavirus.
 XX
 PN US5744142-A.
 XX
 PD 28-APR-1998.
 XX
 PF 07-JUN-1995; 95US-00475782.
 XX
 PR 03-SEP-1992; 92US-00941371.
 PR 16-MAR-1993; 93US-00032869.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Schiller JT, Kirnbauer R, Lowy DR;
 XX
 DR WPI; 1998-271010/24.
 DR N-PSDB; AAV25817.
 XX
 PT Immunisation of mammals and humans against papillomavirus infection -
 PT comprises administering recombinant self-assembled capsid proteins
 PT containing conformational epitopes.
 XX
 PS Claim 1; Col 21-24; 20pp; English.
 XX
 CC The present sequence represents the human papillomavirus HPV 16 L1
 CC conformational epitope, from the present invention. The present invention
 CC describes the immunisation of a host mammal against a papillomavirus
 CC comprising administering to the host, according to an immunising
 CC schedule, a self-assembled Human Papilloma Virus (HPV16) capsid (I)
 CC containing at least 1 papillomavirus L1 conformational epitope. The
 CC papillomavirus L1 conformational epitopes are produced by letting a
 CC genetic construct comprising a papillomavirus L1 gene direct recombinant
 CC expression of the conformational epitope in a transformed eukaryotic host
 CC cell by self-assembly of papillomavirus capsids containing a L1
 CC polypeptide having an amino acid sequence encoded by the nucleic acid
 CC sequence given in the specification (see AAV25817). (I) are useful for
 CC the diagnosis of and as vaccines for the prevention of papillomavirus
 CC infections and their benign and malignant sequelae in vertebrates.
 CC Recombinant conformed papillomavirus proteins are provided which can be

CC used to produce renewable papillomavirus reagents of any selected species
 CC and type in cell culture. The self-assembled recombinant L1 capsid
 CC protein has the efficacy of intact papillomavirus particles to induce
 CC high levels of neutralising antiserum, in contrast to prior art L1
 CC protein extracted from recombinant bacteria or denatured virions
 XX

SQ Sequence 505 AA;

Query Match 100.0%; Score 49; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0;

Qy 1 AGVDNRECI 9

Db 139 AGVDNRECI 147

RESULT 13

AAW44142

ID AAW44142 standard; protein; 505 AA.

XX

AC AAW44142;

XX 05-MAY-1998 (first entry);

XX Human papillomavirus wild-type L1 protein.

XX Human; papillomavirus; bovine; L1 conformational protein; antibody;
 KW detection; diagnosis.

XX Human papillomavirus.

OS US5709996-A.

XX 20-JAN-1998.

XX 07-JUN-1995; 95US-00472673.

XX 03-SEP-1992; 92US-00941371.

XX 16-MAR-1993; 93US-00032869.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Schiller JT, Kirnbauer R, Lowy DR;

XX WPI; 1998-109811/10.

XX N-PSDB; AAV12161.

XX Assay for papilloma virus - using antibody directed against recombinant
 PT L1 polypeptide.
 XX Claim 1; Col 21-24; 34pp; English.

XX An method has been developed for detecting papilloma virus in a specimen
 CC from a mammal. The method comprises contacting the sample with an
 CC antibody and detecting any antibody binding. The antibody is a polyclonal
 CC or monoclonal antibody produced by transforming a eukaryotic host cell
 CC with a genetic construct containing a papillomavirus L1 gene so that the
 CC cell expresses self-assembled papillomavirus-like particles containing at
 CC least 1 papillomavirus L1 conformational epitope. The present sequence
 CC represents wild-type human papillomavirus L1 protein. The method is for
 CC the diagnosis of papillomavirus infections in mammals

XX Sequence 505 AA;

Query Match 100.0%; Score 49; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0;

Qy 1 AGVDNRECI 9

Db 139 AGVDNRECI 147

Qy

Db

RESULT 14

AAW44143

ID AAW44143 standard; protein; 505 AA.

XX

AC AAW44143;

XX 05-MAY-1998 (first entry)

XX Human papillomavirus prototype L1 protein.

DE Human; papillomavirus; bovine; L1 conformational protein; antibody;
 KW detection; diagnosis.

XX Human papillomavirus.

OS US5709996-A.

XX 20-JAN-1998.

XX 07-JUN-1995; 95US-00472673.

XX 03-SEP-1992; 92US-00941371.

XX 16-MAR-1993; 93US-00032869.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Schiller JT, Kirnbauer R, Lowy DR;

XX WPI; 1998-109811/10.

XX N-PSDB; AAV12162.

XX Assay for papilloma virus - using antibody directed against recombinant
 PT L1 polypeptide.
 XX Disclosure; Col 17-20; 34pp; English.

XX An method has been developed for detecting papilloma virus in a specimen
 CC from a mammal. The method comprises contacting the sample with an
 CC antibody and detecting any antibody binding. The antibody is a polyclonal
 CC or monoclonal antibody produced by transforming a eukaryotic host cell
 CC with a genetic construct containing a papillomavirus L1 gene so that the
 CC cell expresses self-assembled papillomavirus-like particles containing at
 CC least 1 papillomavirus L1 conformational epitope. The present sequence
 CC represents prototype human papillomavirus L1 protein. The method is for
 CC the diagnosis of papillomavirus infections in mammals

XX Sequence 505 AA;

Query Match 100.0%; Score 49; DB 2; Length 505;

Best Local Similarity 100.0%; Pred. No. 0.23; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0;

Qy 1 AGVDNRECI 9

Db 139 AGVDNRECI 147

RESULT 15

AAW54034

ID. AAW54034 standard; protein; 505 AA.

XX

AC AAW54034;

XX 17-OCT-2003 (revised)

DT 10-AUG-1998 (first entry)

XX HPV16 L1 gene protein #1.

XX Human papillomavirus 16; HPV17; L1 gene; infection; antibody detection.

XX Human papillomavirus type 16.

PN US5756284-A.
 XX 26-MAY-1998.
 XX
 XX 07-JUN-1995; 95US-00472672.
 XX 03-SEP-1992; 92US-00941371.
 PR 16-MAR-1993; 93US-00032869.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Kirnbauer R, Lowy DR, Schiller JT;
 XX WPI; 1998-321522/28.
 DR N-PSDB; AAV23923.
 XX Assay for antibodies to papilloma virus - using self-assembled
 PT papillomavirus-like particles comprising L1 polypeptide.
 XX Disclosure; Col 17-20; 21pp; English.
 XX This sequence represents the human papillomavirus 16 (HPV16) L1 gene
 CC protein. The protein can be used in the method of the invention for
 CC detecting antibodies to a papilloma virus in a sample from a mammal. The
 CC method comprises: (a) providing self-assembled papillomavirus-like
 CC particles comprising a L1 polypeptide, where the self-assembled
 CC papillomavirus-like particles comprise at least one L1 conformational
 CC epitope and are produced by permitting a genetic construct comprising a
 CC papillomavirus L1 gene to direct recombinant expression in a transformed
 CC eukaryotic host cell; (b) contacting the sample with the self-assembled
 CC papillomavirus-like particles; and (c) detecting any antibody binding to
 CC the self-assembled papillomavirus-like particles. The method can be used
 CC for detecting antibodies to HPV16 as an indication of infection. (Updated
 CC on 17-OCT-2003 to standardise OS field)
 XX SQ Sequence 505 AA;
 Query Match 100.0%; Score 49; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23; 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGVDNRECI 9
 |||||
 Db 139 AGVDNRECI 147
 RESULT 16
 AAW54073
 ID AAW54073 standard; protein; 505 AA.
 XX
 XX AAW54073;
 AC
 XX 17-OCT-2003 (revised)
 DT 10-AUG-1998 (first entry)
 XX HPV16 L1 gene protein #2.
 DE Human papillomavirus 16; HPV17; L1 gene; infection; antibody detection.
 KW Human papillomavirus type 16.
 OS Human papillomavirus type 16.
 XX US5756284-A.
 PN
 XX 26-MAY-1998.
 PD
 XX 07-JUN-1995; 95US-00472672.
 PF
 XX 03-SEP-1992; 92US-00941371.
 PR 16-MAR-1993; 93US-00032869.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Kirnbauer R, Lowy DR, Schiller JT;
 PI

XX WPI; 1998-321522/28.
 DR N-PSDB; AAV24099.
 XX Assay for antibodies to papilloma virus - using self-assembled
 PT papillomavirus-like particles comprising L1 polypeptide.
 XX Claim 1; Col 21-24; 21pp; English.
 XX This sequence represents the human papillomavirus 16 (HPV16) L1 gene
 CC protein. The protein can be used in the method of the invention for
 CC detecting antibodies to a papilloma virus in a sample from a mammal. The
 CC method comprises: (a) providing self-assembled papillomavirus-like
 CC particles comprising a L1 polypeptide, where the self-assembled
 CC papillomavirus-like particles comprise at least one L1 conformational
 CC epitope and are produced by permitting a genetic construct comprising a
 CC papillomavirus L1 gene to direct recombinant expression in a transformed
 CC eukaryotic host cell; (b) contacting the sample with the self-assembled
 CC papillomavirus-like particles; and (c) detecting any antibody binding to
 CC the self-assembled papillomavirus-like particles. The method can be used
 CC for detecting antibodies to HPV16 as an indication of infection. (Updated
 CC on 17-OCT-2003 to standardise OS field)
 XX SQ Sequence 505 AA;
 Query Match 100.0%; Score 49; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23; 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGVDNRECI 9
 |||||
 Db 139 AGVDNRECI 147
 RESULT 17
 AAW96747
 ID AAW96747 standard; protein; 505 AA.
 XX
 XX AAW96747;
 AC
 XX 14-APR-1999 (first entry)
 DT Human papillomavirus (HPV16) L1 major capsid protein fragment.
 DE HPV16; L1 major capsid protein; HPV16 capsid; L1 epitope;
 KW subunit vaccine; HPV infection; wart; laryngeal papillomatosis.
 XX Human papillomavirus.
 OS US5871998-A.
 PN
 XX 16-FEB-1999.
 PD
 XX 07-JUN-1995; 95US-00472678.
 PF
 XX 03-SEP-1992; 92US-00941371.
 PR 16-MAR-1993; 93US-00032869.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Kirnbauer R, Lowy DR, Schiller JT;
 PI
 XX WPI; 1999-166654/14.
 DR N-PSDB; AAX15111.
 XX Self-assembled, recombinant human papilloma virus 16 capsids - for use in
 PT subunit vaccines and as immunoassay reagents.
 XX Disclosure; Col 17-20; 20pp; English.
 XX The present sequence represents a fragment of the prototype Human
 CC papillomavirus (HPV16) L1 major capsid protein. The sequence is used to
 CC construct a HPV16 capsid, containing at least one L1 conformational

CC epitope. The capsids are produced by self-assembly following recombinant
 CC expression of a gene construct, containing an L1 gene, in a transformed
 CC eukaryotic cell. The capsids are used in subunit vaccines to generate a
 CC high-titre neutralizing antibody response in vertebrates, for treatment
 CC and prevention of HPV infection and its benign or malignant consequences,
 CC e.g. warts and laryngeal papillomatosis, including vaccination of
 CC pregnant women for passive protection of infants through the placenta or
 CC milk; as immunoassay reagents for diagnosing HPV infection; as reagents
 CC to detect cellular or humoral immunity, and as immunogens to generate
 CC antibodies (useful as immunoassay reagents for detecting capsid proteins)
 XX
 SQ Sequence 505 AA;

Query Match 100.0%; Score 49; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23; 0; Mismatches 0; Gaps 0;
 Matches 9; Conservative 0;

QY 1 AGVDNRECI 9
 |||||
 DB 139 AGVDNRECI 147

RESULT 18
 AAW96748
 ID AAW96748 standard; protein; 505 AA.

XX AC AAW96748;

XX DT 14-APR-1999 (first entry)

XX DE Human papillomavirus (HPV16) L1 major capsid protein fragment.

XX KW HPV16; L1 major capsid protein; HPV16 capsid; L1 epitope;
 subunit vaccine; HPV infection; wart; laryngeal papillomatosis.

XX OS Human papillomavirus.

XX PN US5871998-A.

XX PD 16-FEB-1999.

XX PF 07-JUN-1995; 95US-00472678.

XX PR 03-SEP-1992; 92US-00941371.

XX PR 16-MAR-1993; 93US-00032869.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Kirnbauer R, Lowy DR, Schiller JT;

XX DR WPI; 1999-166654/14.

XX DR N-PSDB; AAX15112.

XX PT Self-assembled, recombinant human papilloma virus 16 capsids - for use in
 subunit vaccines and as immunoassay reagents.

XX PS Disclosure; Col 19-24; 20pp; English.

XX CC The present sequence represents a fragment of the Human papillomavirus
 CC (HPV16) L1 major capsid protein. The sequence is used to construct a
 CC HPV16 capsid, containing at least one L1 conformational epitope. The
 CC capsids are produced by self-assembly following recombinant expression of
 CC a gene construct, containing an L1 gene, in a transformed eukaryotic
 CC cell. The capsids are used in subunit vaccines to generate a high-titre
 CC neutralizing antibody response in vertebrates, for treatment and
 CC prevention of HPV infection and its benign or malignant consequences,
 CC e.g. warts and laryngeal papillomatosis, including vaccination of
 CC pregnant women for passive protection of infants through the placenta or
 CC milk; as immunoassay reagents for diagnosing HPV infection; as reagents
 CC to detect cellular or humoral immunity, and as immunogens to generate
 CC antibodies (useful as immunoassay reagents for detecting capsid proteins)
 XX

SQ Sequence 505 AA;

Query Match 100.0%; Score 49; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23; 0; Mismatches 0; Gaps 0;
 Matches 9; Conservative 0;

QY 1 AGVDNRECI 9
 |||||
 DB 139 AGVDNRECI 147

RESULT 19

AAAY08019

ID AAY08019 standard; protein; 505 AA.

XX AC AAY08019;

XX DT 27-AUG-2003 (revised)

XX DT 08-JUL-1999 (first entry)

XX DE Human papilloma virus L1 protein.

XX KW L1 protein; capsomer; virus; vaccine; infection; treatment; prevention;
 cervical carcinoma; fusion protein; anti-capsid; antibody; antigenicity.

XX OS Human papillomavirus.

XX PN WO9918220-A1.

XX PD 15-APR-1999.

XX PF 06-OCT-1998; 98WO-US020965.

XX PR 06-OCT-1997; 97US-00944368.

XX PA (LOYO) UNIV LOYOLA CHICAGO.

XX PI Giesmann L, Mueller M;

XX DR WPI; 1999-264026/22.

XX DR N-PSDB; AAX37566.

XX PT Human papilloma virus (HPV) L1 fusion protein capsomers, used in vaccines
 against HPV infection.

XX PS Disclosure; Page 31-32; 48pp; English.

XX CC This invention describes novel vaccines comprising a human papilloma
 CC virus (HPV) capsomer, composed of a HPV L1 fusion protein or a HPV L1
 CC truncated protein. The vaccines can prevent or treat human papilloma
 CC virus infection. Such therapeutic vaccinations can be used for relief
 CC against, e.g. cervical carcinoma. Construction of chimeric proteins
 CC comprising amino acid residues from L1 protein, and e.g. E6 or E7
 CC protein, which give rise to chimeric capsomers, combines prophylactic and
 CC therapeutic functions of a vaccine. Capsomers can promote elimination of
 CC persistently infected cells. Capsomers can also escape neutralization by
 CC pre-existing anti-capsid antibodies and hence possess longer circulating
 CC half-life as compared to chimeric virus-like particles. The fusion
 CC protein, which forms the capsomer, provides increased antigenicity.
 CC (Updated on 27-AUG-2003 to reflect OS field.)

SQ Sequence 505 AA;

Query Match 100.0%; Score 49; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
 |||||
 DB 139 AGVDNRECI 147

RESULT 20

AAAY3580

```

ID AAY53580 standard; protein; 505 AA.
XX
AC AAY53580;
XX
XX
DT 18-FEB-2000 (first entry)
DE HPV16 L1 protein.
XX
KW Human papillomavirus; HPV16; L1; construct; major capsid protein; wart;
KW recombinant; expression; epitope; antibody; capsomer; diagnosis; vaccine;
KW antiserum; communicable papillomatosis; lesion; benign; infection;
KW laryngeal papillomatosis.
XX
XX
OS Human papillomavirus.
XX
XX US5985610-A.
XX
XX 16-NOV-1999.
XX
XX 07-JUN-1995; 95US-00484503.
XX
XX 03-SEP-1992; 92US-00941371.
XX 16-MAR-1993; 93US-00032869.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Kirnbauer R, Schiller JT, Lowy DR;
XX WPI; 2000-012790/01.
XX N-PSDB; AA240564.
XX
XX Genetic construct comprising a human papillomavirus gene is useful in
XX diagnosis and prevention of papillomavirus infections.
XX
XX Disclosure; Col 21-22; 21pp; English.
XX
XX This sequence represents the human papillomavirus type 16 (HPV16) L1
XX protein. The invention relates to a genetic construct comprising an HPV16
XX gene for the L1 major capsid protein, which directs recombinant
XX expression of L1 conformational epitopes used for raising neutralizing
XX antibodies in a host cell by self-assembly of capsomer structures
XX comprising a HPV16 L1 polypeptide. The construct is used for the
XX diagnosis and prevention of papillomavirus infections. The recombinant L1
XX papillomavirus capsid proteins are able to induce high titer neutralizing
XX antiserum making them suitable for use against communicable
XX papillomatosis. The vaccines can be used to treat productive
XX papillomavirus lesions that occur in benign infections such as warts or
XX laryngeal papillomatosis
XX
XX Sequence 505 AA;
XX
Query Match 100.0%; Score 49; DB 3; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
Db 139 AGVDNRECI 147

RESULT 21
AAY53579
ID AAY53579 standard; protein; 505 AA.
XX
XX AAY53579;
XX
XX 18-FEB-2000 (first entry)
XX
XX HPV16 L1 protein - corrected version.
XX
KW Human papillomavirus; HPV16; L1; construct; major capsid protein; wart;
KW recombinant; expression; epitope; antibody; capsomer; diagnosis; vaccine;
KW antiserum; communicable papillomatosis; lesion; benign; infection;

ID AAY53580 standard; protein; 505 AA.
XX
XX
XX
DT 18-FEB-2000 (first entry)
DE HPV16 L1 protein
XX
KW Human papillomavirus; HPV16; L1; construct; major capsid protein; wart;
KW recombinant; expression; epitope; antibody; capsomer; diagnosis; vaccine;
KW laryngeal papillomatosis.
XX
XX
OS Human papillomavirus.
XX
XX US5985610-A.
XX
XX 16-NOV-1999.
XX
XX 07-JUN-1995; 95US-00484503.
XX
XX 03-SEP-1992; 92US-00941371.
XX 16-MAR-1993; 93US-00032869.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Kirnbauer R, Schiller JT, Lowy DR;
XX WPI; 2000-012790/01.
XX N-PSDB; AA240563.
XX
XX Genetic construct comprising a human papillomavirus gene is useful in
XX diagnosis and prevention of papillomavirus infections.
XX
XX Disclosure; Col 17-20; 21pp; English.
XX
XX This sequence represents a corrected version of the human papillomavirus
XX type 16 (HPV16) L1 protein as compared to the HPV16 L1 gene sequence
XX given in the prototype sequence (GenBank Accession No: K02718). The
XX invention relates to a genetic construct comprising an HPV16 gene for the
XX L1 major capsid protein, which directs recombinant expression of L1
XX conformational epitopes used for raising neutralizing antibodies in a
XX host cell by self-assembly of capsomer structures comprising a HPV16 L1
XX polypeptide. The construct is used for the diagnosis and prevention of
XX papillomavirus infections. The recombinant L1 papillomavirus capsid
XX proteins are able to induce high titer neutralizing antiserum making them
XX suitable for use against communicable papillomatosis. The vaccines can be
XX used to treat productive papillomavirus lesions that occur in benign
XX infections such as warts or laryngeal papillomatosis
XX
XX Sequence 505 AA;
XX
Query Match 100.0%; Score 49; DB 3; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
Db 139 AGVDNRECI 147

RESULT 22
AAY57720
ID AAY57720 standard; protein; 505 AA.
XX
XX AAY57720;
XX
XX 14-MAR-2000 (first entry)
XX
XX Human papillomavirus L1 fusion protein SEQ ID NO:2.
XX
KW Human papillomavirus; HPV; L1 fusion protein; vaccine; cytostatic;

```

KW viral capsomere; virucide; dermatological; malignant tumour formation;
 KW cervical cancer; cervical intraepithelial neoplasia; genital wart;
 KW condylomata acuminata.
 XX
 OS Human papillomavirus.
 OS Synthetic.
 XX
 PN CA2229955-A1.
 XX
 PD 20-AUG-1999.
 XX
 PF 20-FEB-1998; 98CA-02229955.
 XX
 PR 20-FEB-1998; 98CA-02229955.
 XX
 PA (MEDI-) MEDIGENE GMBH.
 XX
 PI Burger A, Hallek M;
 XX
 DR WPI; 2000-063092/06.
 DR N-PSDB; AAZ48174.
 XX
 PT Fusion proteins comprising papillomavirus specific proteins useful for
 PT vaccinating against malignant tumors of the anogenital tract such as
 PT cervical carcinomas.
 XX
 PS Claim 5; Page 32-33; 46pp; English.
 XX
 CC The present sequence represents the specifically claimed fusion protein
 CC designated L1, which comprises 2 amino acids sequences from 2 different
 CC papillomavirus specific (PVS) proteins. The fusion protein may be
 CC administered for preventing and treating papillomavirus infections in
 CC humans and animals. Papillomaviruses are implicated in the pathology of
 CC malignant tumour formation in the anogenital tract (of these tumours,
 CC cervical cancer is the most frequent (50000 cases/year) and in the
 CC formation of precursor lesions of cervical intraepithelial neoplasia
 CC (CIN). Papillomaviruses also cause benign genital warts such as
 CC condylomata acuminata. However, the type and severity of disease caused
 CC by the papillomavirus is dependent on the strain causing the infection
 XX
 SQ Sequence 505 AA;
 Query Match 100.0%; Score 49; DB 3; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGVDNRECI 9
 |||||
 DB 139 AGVDNRECI 147
 RESULT 23
 ABB77479
 ID ABB77479 standard; protein; 505 AA.
 XX
 AC ABB77479;
 XX
 DT 22-JUL-2002 (first entry)
 XX
 DE HPV16-L1 2.
 XX
 KW Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;
 KW HPV16-L1; cytostatic; virucide.
 XX
 OS Human papillomavirus.
 XX
 PN WO200238769-A2.
 XX
 PD 16-MAY-2002.
 XX
 PF 19-SEP-2001; 2001WO-DE003618.
 XX
 PR 09-NOV-2000; 2000DE-01055545.
 XX
 CC The invention relates to DNA sequences (I) that encode human papilloma
 CC virus (HPV) 16 L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)
 CC or proteins with the biological activity of L1 and L2. Expression vectors
 CC containing (I) or a similar sequence encoding an L1/E7 fusion protein
 CC (ABL58979-ABL58981) and the proteins encoded by them (ABB77478-ABB77483),
 CC are useful in vaccines, especially for control of cervical cancer. (I)
 CC are also useful for recombinant production of L1 and L2 proteins. (I) are
 CC optimised for codon usage in eukaryotic cells and provide high yields of
 CC L1/L2 or their fusions, without the use of viral vectors
 XX
 SQ Sequence 505 AA;
 Query Match 100.0%; Score 49; DB 5; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGVDNRECI 9
 |||||
 DB 139 AGVDNRECI 147
 RESULT 24
 ABB77478
 ID ABB77478 standard; protein; 505 AA.
 XX
 AC ABB77478;
 XX
 DT 22-JUL-2002 (first entry)
 XX
 DE HPV16-L1 1.
 XX
 KW Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;
 KW HPV16-L1; cytostatic; virucide.
 XX
 OS Human papillomavirus.
 XX
 PN WO200238769-A2.
 XX
 PD 16-MAY-2002.
 XX
 PF 19-SEP-2001; 2001WO-DE003618.
 XX
 PR 09-NOV-2000; 2000DE-01055545.
 XX
 CC (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 CC (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
 XX
 PI Mueller M, Leder C, Kleinschmidt J, Sonnenwald U, Biemelt S;
 XX
 DR WPI; 2002-426950/45.
 DR N-PSDB; ABL58977, ABL58982.
 XX
 PT New DNA sequences encoding human papilloma virus L1 or L2 protein, useful
 PT in vaccines, are optimized for high-level expression in eukaryotic cells.
 XX
 PS Claim 1; Fig 5; 39pp; German.
 XX
 CC The invention relates to DNA sequences (I) that encode human papilloma
 CC virus (HPV) 16 L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)
 CC or proteins with the biological activity of L1 and L2. Expression vectors
 CC containing (I) or a similar sequence encoding an L1/E7 fusion protein
 CC (ABL58979-ABL58981) and the proteins encoded by them (ABB77478-ABB77483),
 CC are useful in vaccines, especially for control of cervical cancer. (I)
 CC are also useful for recombinant production of L1 and L2 proteins. (I) are
 CC optimised for codon usage in eukaryotic cells and provide high yields of
 CC L1/L2 or their fusions, without the use of viral vectors
 XX
 SQ Sequence 505 AA;
 Query Match 100.0%; Score 49; DB 5; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGVDNRECI 9
 |||||
 DB 139 AGVDNRECI 147
 RESULT 25
 ABB77479
 ID ABB77479 standard; protein; 505 AA.
 XX
 AC ABB77479;
 XX
 DT 22-JUL-2002 (first entry)
 XX
 DE HPV16-L1 2.
 XX
 KW Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;
 KW HPV16-L1; cytostatic; virucide.
 XX
 OS Human papillomavirus.
 XX
 PN WO200238769-A2.
 XX
 PD 16-MAY-2002.
 XX
 PF 19-SEP-2001; 2001WO-DE003618.
 XX
 PR 09-NOV-2000; 2000DE-01055545.
 XX
 CC (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 CC (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
 XX
 PI Mueller M, Leder C, Kleinschmidt J, Sonnenwald U, Biemelt S;
 XX
 DR WPI; 2002-426950/45.
 DR N-PSDB; ABL58976, ABL58977.
 XX
 PT New DNA sequences encoding human papilloma virus L1 or L2 protein, useful
 PT in vaccines, are optimized for high-level expression in eukaryotic cells.
 XX
 PS Claim 1; Fig 5; 39pp; German.
 XX
 CC The invention relates to DNA sequences (I) that encode human papilloma
 CC virus (HPV) 16 L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)

XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 PA (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
 XX
 PI Mueller M, Leder C, Kleinschmidt J, Sonnenwald U, Biemelt S;
 XX
 DR WPI; 2002-426950/45.
 DR N-PSDB; ABL58977, ABL58982.
 XX
 PT New DNA sequences encoding human papilloma virus L1 or L2 protein, useful
 PT in vaccines, are optimized for high-level expression in eukaryotic cells.
 XX
 PS Claim 1; Fig 6; 39pp; German.
 XX
 CC The invention relates to DNA sequences (I) that encode human papilloma
 CC virus (HPV) 16 L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)
 CC or proteins with the biological activity of L1 and L2. Expression vectors
 CC containing (I) or a similar sequence encoding an L1/E7 fusion protein
 CC (ABL58979-ABL58981) and the proteins encoded by them (ABB77478-ABB77483),
 CC are useful in vaccines, especially for control of cervical cancer. (I)
 CC are also useful for recombinant production of L1 and L2 proteins. (I) are
 CC optimised for codon usage in eukaryotic cells and provide high yields of
 CC L1/L2 or their fusions, without the use of viral vectors
 XX
 SQ Sequence 505 AA;
 Query Match 100.0%; Score 49; DB 5; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGVDNRECI 9
 |||||
 DB 139 AGVDNRECI 147
 RESULT 24
 ABB77478
 ID ABB77478 standard; protein; 505 AA.
 XX
 AC ABB77478;
 XX
 DT 22-JUL-2002 (first entry)
 XX
 DE HPV16-L1 1.
 XX
 KW Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;
 KW HPV16-L1; cytostatic; virucide.
 XX
 OS Human papillomavirus.
 XX
 PN WO200238769-A2.
 XX
 PD 16-MAY-2002.
 XX
 PF 19-SEP-2001; 2001WO-DE003618.
 XX
 PR 09-NOV-2000; 2000DE-01055545.
 XX
 CC (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 CC (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
 XX
 PI Mueller M, Leder C, Kleinschmidt J, Sonnenwald U, Biemelt S;
 XX
 DR WPI; 2002-426950/45.
 DR N-PSDB; ABL58976.
 XX
 PT New DNA sequences encoding human papilloma virus L1 or L2 protein, useful
 PT in vaccines, are optimized for high-level expression in eukaryotic cells.
 XX
 PS Claim 1; Fig 5; 39pp; German.
 XX
 CC The invention relates to DNA sequences (I) that encode human papilloma
 CC virus (HPV) 16 L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)

CC or proteins with the biological activity of L1 and L2. Expression vectors
 CC containing (I) or a similar sequence encoding an L1/E7 fusion protein
 CC (AB58979-AB58981) and the proteins encoded by them (AB577478-AB577483),
 CC are useful in vaccines, especially for control of cervical cancer. (I)
 CC are also useful for recombinant production of L1 and L2 proteins. (I) are
 CC optimised for codon usage in eukaryotic cells and provide high yields of
 CC L1/L2 or their fusions, without the use of viral vectors

XX Sequence 505 AA;

Query Match 100.0%; Score 49; DB 5; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
 |||||
 Db 139 AGVDNRECI 147

RESULT 25
 ABG75800
 ID ABG75800 standard; protein; 505 AA.

XX AC ABG75800;
 XX 23-OCT-2003 (revised)
 DT 08-MAY-2003 (first entry)
 XX Human papillomavirus 16 L1 protein, clone P114/16/11.

XX HPV16; virus-like particle; VLP; infection; cancer; tumour; L1;
 KW anogenital tract; uterine cervix; clone P114/16/11; virucide; vaccine.

XX Human papillomavirus type 16.

XX US2002168372-A1.
 XX 14-NOV-2002.

XX 29-SEP-1998; 98US-00162904.
 XX 16-JUL-1993; 93US-00092528.
 PR 01-MAY-1996; 96US-00641570.
 PR 27-JUN-1997; 97US-00884168.

XX (DURS/) DURST M.
 PA (GISS/) GISSMANN L.

XX Durst M, Gissmann L;

XX WPI; 2003-275319/27.
 DR N-PSDB; ABX11395.

XX New DNA sequence encoding an L1 protein of a papillomavirus capable of
 PT forming virus-like particles, useful for preparing a vaccine against
 PT papillomavirus infection, particularly human papillomavirus (HPV)16
 PT infection.

XX Disclosure; Fig 2; 17pp; English.

XX The invention discloses a DNA sequence encoding an L1 protein of a human
 CC papillomavirus (HPV16) capable of forming virus-like particles (VLPs).
 CC Papillomaviruses infections have been linked to malignant cancer and
 CC malignant tumours of the anogenital tract, particularly cancer of the
 CC uterine cervix. Also disclosed is an antibody that is specifically
 CC directed against the VLP or the L1 protein, a method for determining anti-
 CC -HPV16 virion antibodies in a sample and prophylaxis of papillomavirus
 CC infections. The DNA sequence is useful for preparing a vaccine against
 CC papillomavirus infection, particularly HPV16 infection. The DNA is used
 CC to produce a VLP which can be used in a diagnostic kit to determine anti-
 CC HPV16 virion antibodies in a sample. The VLP can also be used for the
 CC prophylaxis of papillomavirus infections and the DNA used to produce an
 CC L1 protein. The sequence presented is the HPV16 L1 protein, clone

CC P114/16/11. (Updated on 23-OCT-2003 to standardise OS field)
 XX Sequence 505 AA;

Query Match 100.0%; Score 49; DB 6; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
 |||||
 Db 139 AGVDNRECI 147

RESULT 26
 ABG75799.
 ID ABG75799 standard; protein; 505 AA.

XX AC ABG75799;

XX 23-OCT-2003 (revised)
 DT 08-MAY-2003 (first entry)

XX Human papillomavirus 16 L1 protein, clone P114/16/2.

XX HPV16; virus-like particle; VLP; infection; cancer; tumour; L1;
 KW anogenital tract; uterine cervix; clone P114/16/2; virucide; vaccine.

XX Human papillomavirus type 16.

XX US2002168372-A1.

XX 14-NOV-2002.

XX 29-SEP-1998; 98US-00162904.

XX 16-JUL-1993; 93US-00092528.

PR 01-MAY-1996; 96US-00641570.

PR 27-JUN-1997; 97US-00884168.

XX (DURS/) DURST M.

PA (GISS/) GISSMANN L.

XX Durst M, Gissmann L;

XX WPI; 2003-275319/27.

DR N-PSDB; ABX11395.

XX New DNA sequence encoding an L1 protein of a papillomavirus capable of
 PT forming virus-like particles, useful for preparing a vaccine against
 PT papillomavirus infection, particularly human papillomavirus (HPV)16
 PT infection.

XX Disclosure; Fig 1; 17pp; English.

XX The invention discloses a DNA sequence encoding an L1 protein of a human
 CC papillomavirus (HPV16) capable of forming virus-like particles (VLPs).
 CC Papillomaviruses infections have been linked to malignant cancer and
 CC malignant tumours of the anogenital tract, particularly cancer of the
 CC uterine cervix. Also disclosed is an antibody that is specifically
 CC directed against the VLP or the L1 protein, a method for determining anti-
 CC -HPV16 virion antibodies in a sample and prophylaxis of papillomavirus
 CC infections. The DNA sequence is useful for preparing a vaccine against
 CC papillomavirus infection, particularly HPV16 infection. The DNA is used
 CC to produce a VLP which can be used in a diagnostic kit to determine anti-
 CC HPV16 virion antibodies in a sample. The VLP can also be used for the
 CC prophylaxis of papillomavirus infections and the DNA used to produce an
 CC L1 protein. The sequence presented is the HPV16 L1 protein, clone
 CC P114/16/2. (Updated on 23-OCT-2003 to standardise OS field)

XX Sequence 505 AA;

Query Match 100.0%; Score 49; DB 6; Length 505;
 Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
DB 139 AGVDNRECI 147

RESULT 27
ADA27363
ID ADA27363 standard; protein; 505 AA.
XX AC ADA27363;
XX DT 20-NOV-2003 (first entry)
XX DE HPV-16 L1 protein SEQ ID NO:6.
XX KW cell line; American Type Culture Collection PTA-4047; ATCC-4047;
KW baculoviruses; viral recombinant protein; virus-like particle; vaccine;
KW diagnostic reagent; human papillomavirus type 16; HPV-16; L1.
XX OS Human papillomavirus type 16.
XX FH Key Location/Qualifiers
FT Misc-difference 70 /note= "encoded by GTC"
FT FT
FT Misc-difference 99 /note= "encoded by TAC"
XX PN WO2003068804-A2.
XX PD 21-AUG-2003.
XX PF 14-FEB-2003; 2003WO-US004516.
XX PR 14-FEB-2002; 2002US-0356113P.
PR 14-FEB-2002; 2002US-0356118P.
PR 14-FEB-2002; 2002US-0356119P.
PR 14-FEB-2002; 2002US-0356123P.
PR 14-FEB-2002; 2002US-0356128P.
PR 14-FEB-2002; 2002US-0356133P.
PR 14-FEB-2002; 2002US-0356135P.
PR 14-FEB-2002; 2002US-0356150P.
PR 14-FEB-2002; 2002US-0356151P.
PR 14-FEB-2002; 2002US-0356152P.
PR 14-FEB-2002; 2002US-0356154P.
PR 14-FEB-2002; 2002US-0356156P.
PR 14-FEB-2002; 2002US-0356157P.
PR 14-FEB-2002; 2002US-0356161P.
XX PA (NOVA-) NOVAVAX INC.
XX PI Robinson RA;
XX WPI: 2003-646475/61.
DR N-PSDB; ADA27371.
XX New insect cell line designated ATCC PTA-4047, useful for replicating
PT baculoviruses to produce large amounts of recombinant proteins of
PT medical, pharmaceutical and veterinary importance.
XX Disclosure; Page 55-56; 63pp; English.

CC The present invention describes a cell line comprising a cell that is a
CC clone, derivative, mutant and/or transfectant of a cell line designated
CC American Type Culture Collection (ATCC) PTA-4047. The cell upon culture
CC grows continuously and retains the identifying characteristics of the
CC cell line designated ATCC-4047. Also described is a process of making a
CC cell line. The insect cell line is useful in replicating baculoviruses,
CC as a host substrate for baculovirus plaque assays, as a source of insect
CC proteins, acts as a depot for cell transfection to produce recombinant
CC baculoviruses, and in expressing viral recombinant proteins.

CC Extracellular and intracellular viral recombinant proteins and virus-like
CC particles expressed from the cell line are useful as pharmaceutical
CC compositions, vaccines or diagnostic reagents. The present sequence
CC represents a human papillomavirus type 16 (HPV-16) L1 protein, which is
CC used in the exemplification of the present invention.

XX SQ Sequence 505 AA;
Query Match 100.0%; Score 49; DB 6; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.23; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1 AGVDNRECI 9
DB 139 AGVDNRECI 147

RESULT 28
ADA92541
ID ADA92541 standard; protein; 505 AA.
XX AC ADA92541;
XX DT 20-NOV-2003 (first entry)
XX DE HPV-16 L1 codon optimised amino acid sequence SEQ ID NO:6.
XX KW codon optimised; viral capsid protein; virus-like particle; VLP;
KW antigenic; human papillomavirus infection; virucide; vaccine;
KW gene therapy; human papillomavirus type 16; dysplasia; infection; HPV-16;
XX L1.
XX OS Synthetic.
XX OS Human papillomavirus type 16.
XX FH Key Location/Qualifiers
FT Misc-difference 70 /note= "encoded by GTC"
FT FT
FT Misc-difference 99 /note= "encoded by TAC"
XX PN WO2003068933-A2.
XX PD 21-AUG-2003.
XX PF 14-FEB-2003; 2003WO-US004480.
XX PR 14-FEB-2002; 2002US-0356113P.
PR 14-FEB-2002; 2002US-0356118P.
PR 14-FEB-2002; 2002US-0356119P.
PR 14-FEB-2002; 2002US-0356123P.
PR 14-FEB-2002; 2002US-0356128P.
PR 14-FEB-2002; 2002US-0356133P.
PR 14-FEB-2002; 2002US-0356135P.
PR 14-FEB-2002; 2002US-0356150P.
PR 14-FEB-2002; 2002US-0356151P.
PR 14-FEB-2002; 2002US-0356152P.
PR 14-FEB-2002; 2002US-0356154P.
PR 14-FEB-2002; 2002US-0356156P.
PR 14-FEB-2002; 2002US-0356157P.
PR 14-FEB-2002; 2002US-0356161P.
XX PA (NOVA-) NOVAVAX INC.
XX PI Robinson RA;
XX WPI: 2003-689664/65.
DR N-PSDB; ADA92549.
XX New codon optimized polynucleotide encoding a viral capsid protein that
PT self assembles into a virus-like particle, useful for diagnosing,
PT preventing or treating human papillomavirus infections or associated

PT disorders.

XX Disclosure; Page 115-116; 123pp; English.

XX

XX The present invention describes a codon optimised polynucleotide encoding

CC a viral capsid protein that self assembles into a virus-like particle

CC (VLP) that exhibits conformational antigenic epitopes capable of raising

CC neutralising antibodies, where the VLP is expressed from a host cell

CC extracellularly. Also described: (1) a vector comprising the above codon

CC optimised polynucleotide operably linked to a eukaryotic or prokaryotic

CC regulatory control element, capable of replication in prokaryotic and/or

CC eukaryotic host; (2) a host cell comprising the vector; (3) a

CC pharmaceutical or vaccine composition for treating, ameliorating or

CC preventing a papillomavirus related disease or disorder, comprising a

CC multiplicity of VLPs that exhibit conformational antigenic epitopes, and

CC a carrier, diluent or adjuvant; (4) a diagnostic kit for detecting a

CC papillomavirus infection, comprising a multiplicity of VLPs that exhibit

CC conformational antigenic epitopes, and a detection agent comprising a

CC detectable label; (5) a method for preparing the above codon optimised

CC polynucleotide, comprising replacing codons that are underutilised in

CC insect cells with codons that are utilised at high levels in insect

CC cells, to create an initially-modified nucleotide sequence, and modifying

CC the initially-modified nucleotide sequence by choosing a preferred codon

CC for the initially-modified sequence, where the ratio of GC nucleotide

CC pairs to AT nucleotide pairs in the further-modified nucleotide sequence

CC trends towards about 1:1, where the number of palindromic and stem-loop

CC DNA structures in the further-modified nucleotide sequence is minimised,

CC and where the number of transcription and post-transcription repressor

CC elements are minimised; and (6) methods for treating, ameliorating or

CC preventing a papillomavirus related disease or disorder, or for

CC protecting an individual against a papillomavirus infection, comprising

CC administering to an individual an amount of the composition or vaccine

CC cited above. The VLP has virucide activity and can be used in vaccines

CC and in gene therapy. The composition and methods of the present invention

CC are useful in diagnosing, preventing or treating human papillomavirus

CC infections or associated disorders, such as dysplasia. The present

CC sequence represents an HPV-16 codon optimised L1 amino acid sequence from

CC the present invention.

XX

XX Sequence 505 AA;

XX

Query Match 100.0%; Score 49; DB 7; Length 505;

Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9

Db 139 AGVDNRECI 147

RESULT 29

ADA14290

ID ADA14290 standard; protein; 505 AA.

XX

AC ADA14290;

XX

DT 06-NOV-2003 (first entry)

XX

DE HPV-16 L1 protein SEQ ID NO:6.

XX

KW purification; recombinant extracellular virus-like particle;

KW recombinant intracellular virus-like particle; virus-like particle; VLP;

KW virucide; vaccine; Gene therapy; human papillomavirus; HPV; infection;

KW dysplasia; HPV-16; L1.

XX

OS Human papillomavirus type 16.

XX

XX

Key Location/Qualifiers

FT Misc-difference 70

FT /note= "encoded by GTC"

FT Misc-difference 99

FT /note= "encoded by TAC"

XX

XX

PN WO2003068993-A1.

XX

PD 21-AUG-2003.

XX

XX 14-FEB-2003; 2003WO-US004474.

XX

XX 14-FEB-2002; 2002US-0356113P.

XX

XX 14-FEB-2002; 2002US-0356118P.

XX

XX 14-FEB-2002; 2002US-0356119P.

XX

XX 14-FEB-2002; 2002US-0356123P.

XX

XX 14-FEB-2002; 2002US-0356126P.

XX

XX 14-FEB-2002; 2002US-0356133P.

XX

XX 14-FEB-2002; 2002US-0356135P.

XX

XX 14-FEB-2002; 2002US-0356150P.

XX

XX 14-FEB-2002; 2002US-0356151P.

XX

XX 14-FEB-2002; 2002US-0356152P.

XX

XX 14-FEB-2002; 2002US-0356154P.

XX

XX 14-FEB-2002; 2002US-0356156P.

XX

XX 14-FEB-2002; 2002US-0356157P.

XX

XX 14-FEB-2002; 2002US-0356161P.

XX

XX 14-FEB-2002; 2002US-0356162P.

XX

XX (NOVA-) NOVAVAX INC.

XX

XX Robinson RA, Thompson MW;

XX

XX WPI; 2003-679645/64.

XX

XX N-PSDB; ADA14298.

XX

XX Purifying a recombinant human papillomavirus (HPV) L1, useful for

XX diagnosing, preventing or treating HPV infections, comprises clarifying,

XX concentrating and diafiltering cells containing HPV particles.

XX

XX Disclosure; Page 99-100; 111pp; English.

XX

XX The present invention describes a method for purifying a recombinant

XX extracellular or intracellular virus-like particle (VLP). The method

XX comprises harvesting a cell suspension comprising cells containing a

XX plurality of VLPs to produce a harvested supernatant, optionally

XX disrupting the harvested cells to produce cell lysates containing the

XX VLP, clarifying the harvested supernatant, concentrating the clarified

XX supernatant, diafiltering the concentrated supernatant, and recovering

XX the purified recombinant VLP. Also described: (1) a cell line designated

XX as Sf-98 deposited as American Type Culture Collection (ATCC) PTA-4047;

XX (2) producing the cell line described above; (3) host cells that express

XX one or more recombinant gene products with an enhanced yield; (4)

XX producing a foreign protein in an insect cell; (5) nucleic acid sequences

XX that correspond to and code for human papillomavirus (HPV) polypeptides;

XX and (6) pharmaceutical compositions comprising an amount of the

XX recombinant viral gene products, VLPs, agonists, antagonists, or the

XX active fragment of a viral gene product. The VLPs have virucide activity,

XX and can be used in vaccines and in gene therapy. The method is useful in

XX isolating and purifying expressed viral gene products, including VLPs, in

XX vitro. The gene products or particles may be used in detecting,

XX preventing or treating HPV infections and associated symptoms like

XX dysplasia. The present sequence represents an HPV-16 L1 protein sequence,

XX which is given in the exemplification of the present invention.

XX

XX Sequence 505 AA;

XX

Query Match 100.0%; Score 49; DB 7; Length 505;

Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9

Db 139 AGVDNRECI 147

RESULT 30

AAE38615

ID AAE38615 standard; protein; 505 AA.

XX

XX The present invention describes an isolated prepared human papillomavirus
 CC (HPV) epitope (I). (I) has antiviral activity, and can be used in vaccine
 CC production. Peptides and corresponding nucleic acid compositions from the
 CC present invention are useful for stimulating an immune response to HPV by
 CC stimulating the production of CTL or HTL responses, specifically in the
 CC treatment or prophylaxis of HPV infection, in persons who have not
 CC manifested symptoms e.g. genital warts or neoplastic growth. The peptides
 CC can also be used in a tetramer staining assay to assess peripheral blood
 CC mononuclear cells for the presence of antigen-specific CTLs following
 CC exposure to a pathogen or immunogen, and as reagents to evaluate immune
 CC recall responses or evaluate the efficacy of a vaccine. The vaccine
 CC compositions are useful for removing warts or treating HPV infections.
 CC The epitopes for inclusion in an epitope-base vaccine may be selected
 CC from conserved regions of viral or tumour-associated antigens, which
 CC reduces the likelihood of escape mutants, also immunosuppressive epitopes
 CC that may be present in whole antigens can be avoided with the use of
 CC epitope-base vaccines. An additional advantage is the ability to combine
 CC selected epitopes (CTL and HTL) and to modify the composition of the
 CC epitopes achieving enhanced immunogenicity, the major benefit of the
 CC vaccine is that is safe and efficacious. AAB98391 to AAB98477 represent
 CC polypeptide sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 531 AA;

Query Match 100.0%; Score 49; DB 4; Length 531;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGVDNRECI 9
 |||||
 Db 165 AGVDNRECI 173

RESULT 35
 ABB77482
 ID ABB77482 standard; protein; 533 AA.
 XX
 AC ABB77482;
 XX
 DT 22-JUL-2002 (first entry)
 XX
 DE HPV16-L1/E7 fusion protein (short).
 XX
 KW Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;
 KW HPV16-L1; cytostatic; virucide; HPV16-E7; HPV16-L1/E7.
 XX
 OS Human papillomavirus.
 XX
 FH Key Location/Qualifiers
 FT Protein 1..533
 FT /product= "HPV16-L1/E7 fusion protein"
 FT Region 1..473
 FT /product= "HPV16-L1"
 FT Region 474..533
 FT /product= "HPV16-E7"
 XX
 FN WO200238769-A2.
 XX
 PD 16-MAY-2002.
 XX
 PF 19-SEP-2001; 2001WO-DE003618.
 XX
 PR 09-NOV-2000; 2000DE-01055545.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 PA (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
 XX
 PI Mueller M, Leder C, Kleinschmidt J, Sonnwald U, Biemelt S;
 XX
 DR WPI; 2002-426950/45.
 DR N-PSDB; ABL58980.

XX New DNA sequences encoding human papilloma virus L1 or L2 protein, useful
 FT in vaccines, are optimized for high-level expression in eukaryotic cells.
 XX
 PS Claim 4; Fig 9; 39pp; German.
 XX
 CC The invention relates to DNA sequences (I) that encode human papilloma
 CC virus (HPV) 16 L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)
 CC or proteins with the biological activity of L1 and L2. Expression vectors
 CC containing (I) or a similar sequence encoding an L1/E7 fusion protein
 CC (ABL58979-ABL58981) and the proteins encoded by them (ABB77478-ABB77483),
 CC are useful in vaccines, especially for control of cervical cancer. (I)
 CC are also useful for recombinant production of L1 and L2 proteins. (I) are
 CC optimised for codon usage in eukaryotic cells and provide high yields of
 CC L1/L2 or their fusions, without the use of viral vectors
 XX
 SQ Sequence 533 AA;
 Query Match 100.0%; Score 49; DB 5; Length 533;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AGVDNRECI 9
 |||||
 Db 139 AGVDNRECI 147

RESULT 36
 ABB77481
 ID ABB77481 standard; protein; 533 AA.
 XX
 AC ABB77481;
 XX
 DT 22-JUL-2002 (first entry)
 XX
 DE HPV16-L1/E7 fusion protein 1-60.
 XX
 KW Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;
 KW HPV16-L1; cytostatic; virucide; HPV16-E7; HPV16-L1/E7.
 XX
 OS Human papillomavirus.
 XX
 FH Key Location/Qualifiers
 FT Protein 1..533
 FT /product= "HPV16-L1/E7 fusion protein"
 FT Region 1..473
 FT /product= "HPV16-L1"
 FT Region 474..533
 FT /product= "HPV16-E7"
 XX
 FN WO200238769-A2.
 XX
 PD 16-MAY-2002.
 XX
 PF 19-SEP-2001; 2001WO-DE003618.
 XX
 PR 09-NOV-2000; 2000DE-01055545.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 PA (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
 XX
 PI Mueller M, Leder C, Kleinschmidt J, Sonnwald U, Biemelt S;
 XX
 DR WPI; 2002-426950/45.
 DR N-PSDB; ABL58979.
 XX
 PT New DNA sequences encoding human papilloma virus L1 or L2 protein, useful
 FT in vaccines, are optimized for high-level expression in eukaryotic cells.
 XX
 PS Claim 4; Fig 8; 39pp; German.
 XX
 CC The invention relates to DNA sequences (I) that encode human papilloma
 CC virus (HPV) 16 L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)

CC or proteins with the biological activity of L1 and L2. Expression vectors
 CC containing (I) or a similar sequence encoding an L1/E7 fusion protein
 CC (ABL58979-ABL58981) and the proteins encoded by them (ABB77478-ABB77483),
 CC are useful in vaccines, especially for control of cervical cancer. (I)
 CC are also useful for recombinant production of L1 and L2 proteins. (I) are
 CC optimised for codon usage in eukaryotic cells and provide high yields of
 CC L1/L2 or their fusions, without the use of viral vectors

XX SQ Sequence 533 AA;

Query Match 100.0%; Score 49; DB 5; Length 533;
 Best Local Similarity 100.0%; Pred. No. 0.24; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0

QY 1 AGVDNRECI 9
 DB 139 AGVDNRECI 147

RESULT 37
 ABB77483
 ID ABB77483 standard; protein; 542 AA.

XX AC ABB77483;

XX DT 22-JUL-2002 (first entry)

XX DE HPV16-L1/E7 fusion protein (long).

XX KW Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;
 XX KW HPV16-L1; cytostatic; virucide; HPV16-E7; HPV16-L1/E7.

XX OS Human papillomavirus.

XX FH Key Location/Qualifiers
 XX FT Protein 1..542
 XX FT /product= "HPV16-L1/E7 fusion protein"
 XX FT Region 1..473
 XX FT /product= "HPV16-L1"
 XX FT Region 474..542
 XX FT /product= "HPV16-E7"

XX PN WO200239769-A2.

XX PD 16-MAY-2002.

XX PF 19-SEP-2001; 2001WO-DE003618.

XX PR 09-NOV-2000; 2000DE-01055545.

XX PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX PA (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.

XX PI Mueller M, Leder C, Kleinschmidt J, Sonnwald U, Biemelt S;

XX WPI; 2002-426950/45.

XX DR N-PSDB; ABL58981.

XX PT New DNA sequences encoding human papilloma virus L1 or L2 protein, useful
 XX PT in vaccines, are optimized for high-level expression in eukaryotic cells.

XX PS Claim 4; Fig 10; 39pp; German.

XX CC The invention relates to DNA sequences (I) that encode human papilloma
 XX CC virus (HPV) 16 L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)
 XX CC or proteins with the biological activity of L1 and L2. Expression vectors
 XX CC containing (I) or a similar sequence encoding an L1/E7 fusion protein
 XX CC (ABL58979-ABL58981) and the proteins encoded by them (ABB77478-ABB77483),
 XX CC are useful in vaccines, especially for control of cervical cancer. (I)
 XX CC are also useful for recombinant production of L1 and L2 proteins. (I) are
 XX CC optimised for codon usage in eukaryotic cells and provide high yields of
 XX CC L1/L2 or their fusions, without the use of viral vectors

SQ Sequence 542 AA;

Query Match 100.0%; Score 49; DB 5; Length 542;
 Best Local Similarity 100.0%; Pred. No. 0.25;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGVDNRECI 9
 DB 139 AGVDNRECI 147

RESULT 38

XX AC AAB98436
 ID AAB98436 standard; protein; 504 AA.

XX AC AAB98436;

XX DT 22-AUG-2001 (first entry)

XX DE Human papillomavirus protein HPV31 L1.

XX KW Human papillomavirus; human leukocyte antigen; HLA; immune response; HPV;
 XX KW epitope; T cell; identification; vaccine; infection; genital wart;
 XX KW neoplastic growth; antiviral.

XX OS Human papillomavirus.

XX PN WO200141799-A1.

XX PD 14-JUN-2001.

XX PF 11-DEC-2000; 2000WO-US033549.

XX PR 10-DEC-1999; 99US-0172705P.

XX PR 15-AUG-2000; 2000US-00641528.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Chesnut R, Celis E, Grey HM;
 XX WPI; 2001-381497/40.

XX PT An isolated human papilloma virus (HPV) epitope, useful in vaccines for
 XX PT treating HPV infections.

XX PS Disclosure; Page 23-24; 756pp; English.

XX CC The present invention describes an isolated prepared human papillomavirus
 XX CC (HPV) epitope (I). (I) has antiviral activity, and can be used in vaccine
 XX CC production. Peptides and corresponding nucleic acid compositions from the
 XX CC present invention are useful for stimulating an immune response to HPV by
 XX CC stimulating the production of CTL or HTL responses, specifically in the
 XX CC treatment or prophylaxis of HPV infection, in persons who have not
 XX CC manifested symptoms e.g. genital warts or neoplastic growth. The peptides
 XX CC can also be used in a tetramer staining assay to assess peripheral blood
 XX CC mononuclear cells for the presence of antigen-specific CTLs following
 XX CC exposure to a pathogen or immunogen, and as reagents to evaluate immune
 XX CC recall responses or evaluate the efficacy of a vaccine. The vaccine
 XX CC compositions are useful for removing warts or treating HPV infections.
 XX CC The epitopes for inclusion in an epitope-base vaccine may be selected
 XX CC from conserved regions of viral or tumour-associated antigens, which
 XX CC reduces the likelihood of escape mutants, also immunosuppressive epitopes
 XX CC that may be present in whole antigens can be avoided with the use of
 XX CC epitope-base vaccines. An additional advantage is the ability to combine
 XX CC selected epitopes (CTL and HTL) and to modify the composition of the
 XX CC epitopes achieving enhanced immunogenicity, the major benefit of the
 XX CC vaccine is that is safe and efficacious. AAB98391 to AAB98477 represent
 XX CC polypeptide sequences used in the exemplification of the present
 XX CC invention

XX SQ Sequence 504 AA;

Query Match 83.7%; Score 41; DB 4; Length 504;

Best Local Similarity 87.5%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VDNRECI 9
Db 141 GTDNRECI 148
| | | | |

RESULT 39

AAR04969
ID AAR04969 standard; protein; 20 AA.

XX AC AAR04969;
XX DT 25-MAR-2003 (revised)
XX DT 03-OCT-1990 (first entry)

XX Papilloma virus type 16 L1 peptide no. 12.

XX Papilloma virus; PV type 16; IgA; L1; ELISA: cervical cancer.

XX Synthetic.

OS WO9004790-A.

XX PD 03-MAY-1990.

XX PF 28-OCT-1988; 88SE-00003870.

XX PR 28-OCT-1988; 88SE-00003870.

XX PA (MEDS-) MEDSCAND AB.

XX PI Dillner J, Dillner L;

XX DR WPI; 1990-164122/21.

XX Detecting papilloma virus infections - by identifying specific antibodies
PT against partic. viral proteins or fragments, esp. for rapid diagnosis of
PT cervical cancer.

XX PS Claim 1; Page 30; 57pp; English.

XX The peptide is one of 66 overlapping peptides which together cover the
CC entire sequences of the L1 and L2 proteins of human PV type 16. The
CC peptide was found to be reactive with IGA Abs in the sera of patients
CC with type 16 cervical cancer. It can therefore be used in immunoassays
CC for the diagnosis of PV-associated neoplasia, esp. carcinoma of the cervi
CC uteri, and to assess the risk of development of such a carcinoma. See
CC also AAR04958-R05023. (Updated on 25-MAR-2003 to correct PR field.)

XX SQ Sequence 20 AA;

Query Match 79.6%; Score 39; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.69;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VDNRECI 9
Db 1 VDNRECI 7
| | | | |

RESULT 40

AAR32059
ID AAR32059 standard; protein; 51 AA.

XX AC AAR32059;

XX DT 25-MAR-2003 (revised)

XX DT 07-JUN-1993 (first entry)

XX Human IgG reactive HPV16 L1 epitope 167-217.

KW B; epitope; human; papilloma; virus; HPV-16; L1; antisera; ELISA; A1;
KW antigenic index; flexibility; accessibility; hydrophilicity; mice;
KW immunise; HPV16 capsid; papilloma virus like particle; VLP; diagnosis;
KW vaccine.

XX Synthetic.

OS WO9302184-A1.

XX PN 04-FEB-1993.

XX PD 20-JUL-1992; 92WO-AU000364.

XX PF 19-JUL-1991; 91AU-00007322.

XX PR (UYQU) UNIV QUEENSLAND.

XX PA (CLSC-) CLS LTD.

XX PI Frazer I, Zhou J;

XX DR WPI; 1993-058783/07.

XX Prodn. of papilloma virus-like particles - which contain L1 and L2
PT proteins, useful as vaccine against papilloma virus infections and in
PT diagnosis.

XX PS Claim 39; Table 1; 63pp; English.

XX The sequences given in AAR32054-73 are B epitopes which are derived from
CC the human papilloma virus-16 (HPV-16) L1 protein. These epitopes were
CC isolated by screening antisera from immunised animals against a set of
CC overlapping HPV-16 L1 peptides by ELISA. Possible antigenic regions were
CC identified using an antigenic index (AI) on the basis of chain
CC flexibility, high accessibility and a high degree of hydrophilicity. A
CC region with an AI value of over 1.5 was regarded as a predicted B
CC epitope. Five regions identified in this manner were found to be within
CC the 22 peptides to which major reactivity was seen with antisera from
CC mice immunised with synthetic HPV16 capsids. These epitopes can be used
CC to in the production of papilloma virus like particles (VLPs) for use as
CC diagnostic agents and as components of vaccines for use with papilloma
CC virus infections. (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 51 AA;

Query Match 79.6%; Score 39; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 2;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VDNRECI 9
Db 1 VDNRECI 7
| | | | |

Search completed: March 3, 2004, 10:22:47
Job time : 50.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 10:21:39 ; Search time 13.5 Seconds
(without alignments)
64.128 Million cell updates/sec

Title: US-09-980-064-1

Perfect score: 55

Sequence: 1 AQIFNKPYW 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: Pirl.*

2: Pirl2.*

3: Pirl3.*

4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	502	2	S36526 human
2	55	100.0	503	2	S36514 human
3	55	100.0	504	1	PIWL31 human
4	55	100.0	505	1	PIWL35 human
5	55	100.0	528	2	S36520 human
6	55	100.0	531	1	PIWLHS major capsid prote
7	53	96.4	354	1	PIWLRI human
8	53	96.4	499	1	PIWL13 human
9	53	96.4	499	2	S36531 human
10	53	96.4	500	1	PIWL6 human
11	53	96.4	501	1	PIWL11 human
12	53	96.4	502	1	PIWL42 human
13	53	96.4	502	2	S36549 human
14	53	96.4	508	2	S36508 human
15	53	96.4	531	2	S36537 human
16	53	96.4	534	2	S36583 human
17	53	96.4	504	1	PIWL51 human
18	52	94.5	504	1	PIWL51 human
19	50	90.9	499	1	PIWL33 human
20	50	90.9	505	1	PIWLPR human
21	50	90.9	505	1	PIWL39 human
22	50	90.9	524	1	PIWL58 human
23	50	90.9	529	2	S36578 human
24	50	90.9	539	2	S36566 human
25	50	90.9	568	1	PIWL18 human
26	49	89.1	510	1	S15620 human
27	49	89.1	510	1	S15627 human
28	49	89.1	532	2	S36554 human
29	49	89.1	594	2	S36502 human (altern

ALIGNMENTS

RESULT 1

S36526

L1 protein - human papillomavirus type 35H

C;Species: human papillomavirus type 35H

C;Date: 09-Dec-1993 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999

C;Accession: S36526

R;Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A;Description: Primer-directed sequencing of human papillomavirus types.

A;Reference number: S36469

A;Accession: S36526

A;Molecule type: DNA

A;Residues: 1-502

A;Cross-references: EMBL:X74477; NID:g396997; PIDN:CAA52566.1; PID:g397004

A;Experimental source: strain 35H

C;Superfamily: papillomavirus L1 protein

C;Keywords: late protein

Query Match 100.0%; Score 55; DB 2; Length 502;

Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9

Db 302 AQIFNKPYW 310

RESULT 2

S36514

L1 protein - human papillomavirus type 32

C;Species: human papillomavirus type 32

C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999

C;Accession: S36514; B44889

R;Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A;Description: Primer-directed sequencing of human papillomavirus types.

A;Reference number: S36469

A;Accession: S36514

A;Molecule type: DNA

A;Residues: 1-503

A;Cross-references: EMBL:X74475; NID:g396981; PIDN:CAA52554.1; PID:g396988

R;van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnemakers, H.F.; Delius, H.

J. Clin. Microbiol. 30, 1716-1721, 1992

A;Title: General primer polymerase chain reaction in combination with sequence analysis

A;Reference number: A44889; MUID:92332706; PMID:1321168

A;Accession: B44889

A;Molecule type: DNA

A;Residues: 337-368 <VAN>

A;Cross-references: GB:S40336; NID:g251684; PIDN:AAB2563.1; PID:g251685

A;Experimental source: mucosotropic type 32, cervical smear

A;Note: sequence extracted from NCBI backbone (NCBIN:109392, NCBIP:109405)

C;Superfamily: papillomavirus L1 protein

C;Keywords: late protein

Query Match 100.0%; Score 55; DB 2; Length 503;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9

DB 304 AQIFNKPYW 312

RESULT 3

PWL31

L1 protein - human papillomavirus type 31

C;Species: human papillomavirus type 31

A;Note: host Homo sapiens (man)

C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999

C;Accession: G32444

R;Goldborough, M.D.; DiSilvestre, D.; Temple, G.P.; Lorincz, A.T.

Virolgy 171, 306-311, 1989

A;Title: Nucleotide sequence of human papillomavirus type 31: a cervical neoplasia-associated

A;Reference number: A94398; MUID:89299478; PMID:2545036

A;Accession: G32444

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-504 <GOL>

A;Cross-references: GB:J04353; NID:G333048; PIDN:AAA46956.1; PID:G459922

C;Superfamily: papillomavirus L1 protein

C;Keywords: late protein

Query Match 100.0%; Score 55; DB 1; Length 504;

Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9

DB 305 AQIFNKPYW 313

RESULT 4

PWL35

L1 protein - human papillomavirus type 35

C;Species: human papillomavirus type 35

A;Note: host Homo sapiens (man)

C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jul-1999

C;Accession: G40824; C44889

R;Marich, J.E.; Pontsler, A.V.; Rice, S.M.; McGraw, K.A.; Dubensky, T.W.

Virolgy 186, 770-776, 1992

A;Title: The phylogenetic relationship and complete nucleotide sequence of human papillomavirus type 35

A;Reference number: A40824; MUID:92124753; PMID:1310198

A;Accession: G40824

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-505 <MAR>

A;Cross-references: GB:M74117; NID:G333050; PIDN:AAA46972.1; PID:G333058

R;van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnemakers, H.F.; Delius, H.

J. Clin. Microbiol. 30, 1716-1721, 1992

A;Title: General primer polymerase chain reaction in combination with sequence analysis

A;Reference number: A44889; MUID:92332706; PMID:1321168

A;Accession: C44889

A;Molecule type: DNA

A;Residues: 338-370 <VAN>

A;Cross-references: GB:SA0240; NID:G251686; PIDN:AA922564.1; PID:G251687

A;Experimental source: mucosotropic type 35, cervical smear

A;Note: sequence extracted from NCBI backbone (NCBIN:109393, NCBI:109408)

C;Superfamily: papillomavirus L1 protein

C;Keywords: late protein

Query Match 100.0%; Score 55; DB 1; Length 505;

Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9

DB 305 AQIFNKPYW 313

RESULT 5

S36520

L1 protein - human papillomavirus type 34

C;Species: human papillomavirus type 34

C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999

C;Accession: S36520

R;Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A;Description: Primer-directed sequencing of human papillomavirus types.

A;Reference number: S36469

A;Accession: S36520

A;Molecule type: DNA

A;Residues: 1-528 <DEB>

A;Cross-references: EMBL:X74476; NID:G396989; PIDN:CAA52560.1; PID:G396996

C;Superfamily: papillomavirus L1 protein

C;Keywords: late protein

Query Match 100.0%; Score 55; DB 2; Length 528;

Best Local Similarity 100.0%; Pred. No. 0.016;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9

DB 327 AQIFNKPYW 335

RESULT 6

PWLHS

major capsid protein L1 - human papillomavirus type 16

C;Species: human papillomavirus type 16

C;Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 20-Aug-1999

C;Accession: A03640; F44906; T10431

R;Seedorf, K.; Krammer, G.; Durst, M.; Suhai, S.; Rowekamp, W.G.

Virolgy 145, 181-185, 1985

A;Title: Human papillomavirus type 16 DNA sequence.

A;Reference number: A22355; MUID:85246220; PMID:2990099

A;Accession: A03640

A;Molecule type: DNA

A;Residues: 1-531 <SEB>

A;Cross-references: GB:K02718; NID:G333031; PIDN:AAA46943.1; PID:G333037

R;van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnemakers, H.F.; Delius, H.

J. Clin. Microbiol. 30, 1716-1721, 1992

A;Title: General primer polymerase chain reaction in combination with sequence analysis

A;Reference number: A44889; MUID:92332706; PMID:1321168

A;Contents: mucosotropic type 16, cervical smear

A;Accession: F44906

A;Status: not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 363-395 <VAN>

A;Note: sequence extracted from NCBI backbone (NCBIP:109407)

R;Kennedy, I.M.; Haddow, J.K.; Clements, J.B.

J. Virol. 65, 2093-2097, 1991

A;Title: A negative element in the human papillomavirus type 16 genome acts at the level

A;Reference number: Z17014; MUID:91162763; PMID:1848319

A;Accession: T10431

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-531 <KEN>

A;Cross-references: EMBL:K02718; NID:G333031; PIDN:AAA46943.1; PID:G333037

C;Genetics:

A;Gene: L1

C;Superfamily: papillomavirus L1 protein

C;Keywords: late protein

Query Match 100.0%; Score 55; DB 1; Length 531;

Best Local Similarity 100.0%; Pred. No. 0.016;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9


```

Db          330 AQLFNKPYW 338
|||||
RESULT 7
P1WL1
L1 protein - rhesus papillomavirus (type 1)
C;Species: rhesus papillomavirus
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 27-Jan-1995
C;Accession: H38503
R;Ostrow, R.S.; LaBresh, K.V.; Faras, A.J.
Virology 181, 424-429, 1991
A;Title: Characterization of the complete RHPV 1 genomic sequence and an integration loc
A;Reference number: A38503; MUID:91135018; PMID:1847267
A;Accession: H38503
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-354 <OST>
A;Cross-references: EMBL:M37717
A;Superfamily: papillomavirus L1 protein
C;Keywords: late protein

Query Match          96.4%; Score 53; DB 1; Length 354;
Best Local Similarity 88.9%; Pred. No. 0.024;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQLFNKPYW 9
||:|||||
Db 154 AQLFNKPYW 162

RESULT 8
P1WL13
L1 protein - human papillomavirus type 13
C;Species: human papillomavirus type 13
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 16-Jul-1999
C;Accession: H42955; A44889
R;van Rans, A.; Fuse, A.; Fiten, P.; Beuken, E.; Pfister, H.; Burk, R.D.; Opdenakker, G
Virology 190, 587-596, 1992
A;Title: Human papillomavirus type 13 and pygmy chimpanzee papillomavirus type 1: Compar
A;Reference number: A42955; MUID:92391075; PMID:1325697
A;Accession: H42955
A;Molecule type: DNA
A;Residues: 1-499 <VAN>
A;Cross-references: EMBL:X62843; NID:G60295; PIDN:CAA44654.1; PID:G60303
R;van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnemakers, H.F.; Delius, H.
J. Clin. Microbiol. 30, 1716-1721, 1992
A;Title: General primer polymerase chain reaction in combination with sequence analysis
A;Reference number: A44889; MUID:92332706; PMID:1321168
A;Accession: A44889
A;Molecule type: DNA
A;Residues: 331-364 <VA2>
A;Cross-references: GB:S40235; NID:G251682; PIDN:AAB22562.1; PID:G251683
A;Experimental source: mucosotropic type 13, cervical smear
A;Note: sequence extracted from NCBI backbone (NCBIN:109389, NCBIP:109410)
C;Superfamily: papillomavirus L1 protein
C;Keywords: late protein

Query Match          96.4%; Score 53; DB 1; Length 499;
Best Local Similarity 88.9%; Pred. No. 0.035;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQLFNKPYW 9
||:|||||
Db 298 AQLFNKPYW 306

RESULT 9
S36531
L1 protein - human papillomavirus type 53
C;Species: human papillomavirus type 53
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999

```

```

C;Accession: S36531
R;Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A;Description: Primer-directed sequencing of human papillomavirus types.
A;Reference number: S36469
A;Accession: S36531
A;Molecule type: DNA
A;Residues: 1-499 <DEL>
A;Cross-references: EMBL:X74482; NID:G397046; PIDN:CAA52595.1; PID:G397052
C;Superfamily: papillomavirus L1 protein
C;Keywords: late protein

Query Match          96.4%; Score 53; DB 2; Length 499;
Best Local Similarity 88.9%; Pred. No. 0.035;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQLFNKPYW 9
||:|||||
Db 302 AQLFNKPYW 310

RESULT 10
P1WL6
L1 protein - human papillomavirus type 6b
C;Species: human papillomavirus type 6b
C;Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 16-Jul-1999
C;Accession: A03638
R;Schwarz, E.; Durst, M.; Demankowski, C.; Lattermann, O.; Zech, R.; Wolfspurger, E.; Sul
EMBO J. 2, 2341-2348, 1983
A;Title: DNA sequence and genome organization of genital human papillomavirus type 6b.
A;Reference number: A90975; MUID:84131949; PMID:6321162
A;Accession: A03638
A;Molecule type: DNA
A;Residues: 1-500 <SCH>
A;Cross-references: GB:X00203; NID:G60955; PIDN:CAA25026.1; PID:G60964
C;Superfamily: papillomavirus L1 protein
C;Keywords: late protein

Query Match          96.4%; Score 53; DB 1; Length 500;
Best Local Similarity 88.9%; Pred. No. 0.035;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQLFNKPYW 9
||:|||||
Db 300 AQLFNKPYW 308

RESULT 11
P1WL11
L1 protein - human papillomavirus type 11
C;Species: human papillomavirus type 11
C;Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 16-Jul-1999
C;Accession: A03639
R;Dartmann, K.; Schwarz, E.; Gissmann, L.; zur Hausen, H.
Virology 151, 124-130, 1986
A;Title: The nucleotide sequence and genome organization of human papilloma virus type 1
A;Reference number: A94338; MUID:86181601; PMID:3008427
A;Accession: A03639
A;Molecule type: DNA
A;Residues: 1-501 <DAR>
A;Cross-references: GB:M14119; NID:G333026; PIDN:AAA46935.1; PID:G496201
C;Superfamily: papillomavirus L1 protein
C;Keywords: late protein

Query Match          96.4%; Score 53; DB 1; Length 501;
Best Local Similarity 88.9%; Pred. No. 0.035;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQLFNKPYW 9
||:|||||
Db 301 AQLFNKPYW 309

```

```

RESULT 12
P1WL41
L1 protein - pygmy chimpanzee papillomavirus (type 1)
C:Species: pygmy chimpanzee papillomavirus
C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 16-Jul-1999
C:Accession: H36818
R:van Ranst, M.; Fuse, A.; Fiten, P.; Beuken, E.; Pfister, H.; Burk, R.D.; Opdenakker, C.
Virology 190, 587-596, 1992
A:Title: Human papillomavirus type 13 and pygmy chimpanzee papillomavirus type 1: Comparison
A:Reference number: A42955; MUID:92391075; PMID:1325697
A:Accession: H36818
A:Molecule type: DNA
A:Residues: 1-502 <DEL>
A:Cross-references: EMBL:X62844; NID:961010; PIDN:CAA44662.1; PID:961018
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match          96.4%; Score 53; DB 1; Length 502;
Best Local Similarity 88.9%; Pred. No. 0.035;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
||:|||||
Db 301 AQLFNKPYW 309

RESULT 13
P1WL42
L1 protein - human papillomavirus type 42
C:Species: human papillomavirus type 42
A:Note: host Homo sapiens (man)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 27-Jan-1995
C:Accession: G39451
R:Philipp, W.; Honore, N.; Sapp, M.; Cole, S.T.; Strebeck, R.E.
Virology 186, 331-334, 1992
A:Title: Human papillomavirus type 42: new sequence, conserved genome organization.
A:Reference number: A39451; MUID:92087479; PMID:1309278
A:Accession: G39451
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-502 <PHI>
A:Cross-references: GB:M7236
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match          96.4%; Score 53; DB 1; Length 502;
Best Local Similarity 88.9%; Pred. No. 0.035;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
||:|||||
Db 304 AQLFNKPYW 312

RESULT 14
S36549
L1 protein - human papillomavirus type 26
C:Species: human papillomavirus type 26
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36549
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36549
A:Molecule type: DNA
A:Residues: 1-503 <DEL>
A:Cross-references: EMBL:X74472; NID:9396956; PIDN:CAA52535.1; PID:9396963
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match          96.4%; Score 53; DB 2; Length 503;
Best Local Similarity 88.9%; Pred. No. 0.035;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
||:|||||
Db 304 AQLFNKPYW 312

RESULT 15
S36508
L1 protein - human papillomavirus type 30
C:Species: human papillomavirus type 30
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36508
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36508
A:Molecule type: DNA
A:Residues: 1-508 <DEL>
A:Cross-references: EMBL:X74474; NID:9396973; PIDN:CAA52548.1; PID:9396980
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match          96.4%; Score 53; DB 2; Length 508;
Best Local Similarity 88.9%; Pred. No. 0.036;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
||:|||||
Db 311 AQLFNKPYW 319

RESULT 16
S36537
L1 protein - human papillomavirus type 10
C:Species: human papillomavirus type 10
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36537
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36537
A:Molecule type: DNA
A:Residues: 1-531 <DEL>
A:Cross-references: EMBL:X74465; NID:9396901; PIDN:CAA52494.1; PID:9396908
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match          96.4%; Score 53; DB 2; Length 531;
Best Local Similarity 88.9%; Pred. No. 0.037;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
||:|||||
Db 331 AQLFNKPYW 339

RESULT 17
S36583
L1 protein - human papillomavirus type 56
C:Species: human papillomavirus type 56
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36583; H44889
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36583
A:Molecule type: DNA
A:Residues: 1-534 <DEL>
A:Cross-references: EMBL:X74483; NID:9397053; PIDN:CAA52600.1; PID:9397059

```

R;van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnemakers, H.F.; Delius, H.
 J. Clin. Microbiol. 30, 1716-1721, 1992
 A:Title: General primer polymerase chain reaction in combination with sequence analysis
 A:Reference number: A44889; MUID:92332706; PMID:1321168
 A:Accession: H44889
 A:Molecule type: DNA
 A:Residues: 371-402 <VAN>
 A:Cross-references: GB:S40273; NID:g251696; PIDN:AB22569.1; PID:g251697
 A:Experimental source: mucosotropic type 56, cervical smear
 A>Note: sequence inconsistent with the nucleotide translation
 A>Note: sequence extracted from NCBI backbone (NCBIN:109398, NCBIP:109406)
 C:Superfamily: papillomavirus L1 protein
 C:Keywords: late protein

Query Match 96.4%; Score 53; DB 2; Length 534;
 Best Local Similarity 88.9%; Pred. No. 0.038;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AOIFNKPYW 9
 Db 338 AOLFNKPYW 346

RESULT 18

PIWL51

L1 protein - human papillomavirus type 51
 C:Species: human papillomavirus type 51
 A:Note: host Homo sapiens (man)
 C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C:Accession: G40415; G44889
 R:Lungu, O.; Crum, C.P.; Silverstein, S.J.
 J. Virol. 65, 4216-4225, 1991
 A:Title: Biologic properties and nucleotide sequence analysis of human papillomavirus type 51
 A:Reference number: A40415; MUID:91303675; PMID:1649326
 A:Accession: G40415
 A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-504 <LUN>

A:Cross-references: GB:M62877
 R;van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnemakers, H.F.; Delius, H.
 J. Clin. Microbiol. 30, 1716-1721, 1992
 A:Title: General primer polymerase chain reaction in combination with sequence analysis
 A:Reference number: A44889; MUID:92332706; PMID:1321168
 A:Accession: G44889
 A:Molecule type: DNA
 A:Residues: 337-369 <VAN>
 A:Cross-references: GB:S40272; NID:g251694; PIDN:AB22568.1; PID:g251695
 A:Experimental source: mucosotropic type 51, cervical smear
 A>Note: sequence extracted from NCBI backbone (NCBIN:109397, NCBIP:109409)
 C:Superfamily: papillomavirus L1 protein
 C:Keywords: late protein

Query Match 94.5%; Score 52; DB 1; Length 504;
 Best Local Similarity 88.9%; Pred. No. 0.055;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AOIFNKPYW 9
 Db 304 SOLFNKPYW 312

RESULT 19

PIWL33

L1 protein - human papillomavirus type 33
 C:Species: human papillomavirus type 33
 C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
 C:Accession: A03641
 R;Cole, S.T.; Strebeck, R.E.
 J. Virol. 58, 991-995, 1986
 A:Title: Genome organization and nucleotide sequence of human papillomavirus type 33, with
 A:Reference number: A93020; MUID:86200464; PMID:3009902
 A:Accession: A03641
 A:Molecule type: DNA

A:Residues: 1-499 <COL>
 A:Cross-references: GB:WL2732; NID:g333049; PIDN:AAA46964.1; PID:g463183
 C:Superfamily: papillomavirus L1 protein
 C:Keywords: late protein

Query Match 90.9%; Score 50; DB 1; Length 499;
 Best Local Similarity 77.8%; Pred. No. 0.13;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AOIFNKPYW 9
 Db 304 SOLFNKPYW 312

RESULT 20

PIWLPR

L1 protein - human papillomavirus type ME180 (provirus)
 C:Species: human papillomavirus type ME180
 A:Note: host Homo sapiens (man)
 C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 28-Jul-2000
 C:Accession: B40509
 R;Reuter, S.; Delius, H.; Kahn, T.; Hofmann, B.; zur Hausen, H.; Schwarz, E.
 J. Virol. 65, 5564-5568, 1991
 A:Title: Characterization of a novel human papillomavirus DNA in the cervical carcinoma
 A:Reference number: A40509; MUID:91374616; PMID:1716694
 A:Accession: B40509
 A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-505 <REU>

A:Cross-references: GB:M73258; NID:g184383; PIDN:AAF14010.1; PID:g6478870
 C:Superfamily: papillomavirus L1 protein
 C:Keywords: late protein

Query Match 90.9%; Score 50; DB 1; Length 505;
 Best Local Similarity 77.8%; Pred. No. 0.13;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AOIFNKPYW 9
 Db 304 SOLFNKPYW 312

RESULT 21

PIWL39

L1 protein - human papillomavirus type 39
 C:Species: human papillomavirus type 39
 A:Note: host Homo sapiens (man)
 C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999
 C:Accession: H38502
 R;Volpers, C.; Strebeck, R.E.
 Virology 181, 419-423, 1991
 A:Title: Genome organization and nucleotide sequence of human papillomavirus type 39.
 A:Reference number: A38502; MUID:91135017; PMID:1847266
 A:Accession: H38502
 A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-505 <VOL>
 A:Cross-references: GB:M62849; EMBL:M38185; NID:g333245; PIDN:AAA47056.1; PID:g463192
 C:Superfamily: papillomavirus L1 protein
 C:Keywords: late protein

Query Match 90.9%; Score 50; DB 1; Length 505;
 Best Local Similarity 77.8%; Pred. No. 0.13;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AOIFNKPYW 9
 Db 303 SOLFNKPYW 311

RESULT 22

PIWL58

L1 protein - human papillomavirus type 58

C:Species: human papillomavirus type 58

A:Note: host Homo sapiens (man)

C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jul-2000

C:Accession: G36779

R:Kirii, Y.; Iwamoto, S.; Mateukura, T.

Virolgy 185, 424-427, 1991

A:Title: Human papillomavirus type 58 DNA sequence.

A:Reference number: A36779; MUID:92024102; PMID:1656594

A:Accession: G36779

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-524 <KIR>

A:Cross-references: GB:D90400; NID:9222386; PIDN:BA31851.1; PID:G3337104

C:Superfamily: papillomavirus L1 protein

C:Keywords: late protein

Query Match 90.9%; Score 50; DB 1; Length 524;
Best Local Similarity 77.8%; Pred. No. 0.14; 0; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 0;

QY 1 AQIFNKPYW 9

Db 330 SQLEFNKPYW 338

RESULT 23

S36578

L1 protein - human papillomavirus type 52

C:Species: human papillomavirus type 52

C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999

C:Accession: S36578; I44889

R:Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469

A:Accession: S36578

A:Molecule type: DNA

A:Residues: 1-529

A:Cross-references: EMBL:X74481; NID:G937038; PIDN:CAAS2590.1; PID:G937045

R:van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnemakers, H.F.; Delius, H.

J. Clin. Microbiol. 30, 1716-1721, 1992

A:Title: General primer polymerase chain reaction in combination with sequence analysis

A:Reference number: A44889; MUID:92332706; PMID:1321168

A:Accession: I44889

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 368-399 <VAN>

A:Cross-references: GB:S40277; NID:G251698; PIDN:AAB22570.1; PID:G251699

A:Experimental source: type Xa, cervical carcinoma in situ

A:Note: sequence extracted from NCB1 backbone (NCBIN:109399, NCBIP:109447)

C:Superfamily: papillomavirus L1 protein

C:Keywords: late protein

Query Match 90.9%; Score 50; DB 2; Length 529;
Best Local Similarity 77.8%; Pred. No. 0.14; 0; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 0;

QY 1 AQIFNKPYW 9

Db 335 SQLEFNKPYW 343

RESULT 24

S36566

L1 protein - human papillomavirus type 45

C:Species: human papillomavirus type 45

C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999

C:Accession: S36566; F44889

R:Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469

A:Accession: S36566

A:Molecule type: DNA

A:Residues: 1-539

A:Cross-references: EMBL:X74479; NID:G937022; PIDN:CAAS2578.1; PID:G937029

R:van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnemakers, H.F.; Delius, H.

J. Clin. Microbiol. 30, 1716-1721, 1992

A:Title: General primer polymerase chain reaction in combination with sequence analysis

A:Reference number: A44889; MUID:92332706; PMID:1321168

A:Accession: F44889

A:Molecule type: DNA

A:Residues: 366-399 <VAN>

A:Cross-references: GB:S40261; NID:G251692; PIDN:AAB22567.1; PID:G251693

A:Experimental source: mucosotropic type 45, cervical smear

A:Note: sequence extracted from NCB1 backbone (NCBIN:109396, NCBIP:109413)

C:Superfamily: papillomavirus L1 protein

C:Keywords: late protein

Query Match 90.9%; Score 50; DB 2; Length 539;
Best Local Similarity 77.8%; Pred. No. 0.14; 0; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 0;

QY 1 AQIFNKPYW 9

Db 333 SQLEFNKPYW 341

RESULT 25

PIWL18

L1 protein - human papillomavirus type 18

C:Species: human papillomavirus type 18

C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 24-Feb-1994

C:Accession: A26251

R:Coile, S.T.; Danos, O.

J. Mol. Biol. 193, 599-608, 1987

A:Title: Nucleotide sequence and comparative analysis of the human papillomavirus type 1

A:Reference number: A92937; MUID:87283882; PMID:3039146

A:Accession: A26251

A:Molecule type: DNA

A:Residues: 1-568 <COL>

C:Superfamily: papillomavirus L1 protein

C:Keywords: late protein

Query Match 90.9%; Score 50; DB 1; Length 568;
Best Local Similarity 77.8%; Pred. No. 0.15; 0; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 0;

QY 1 AQIFNKPYW 9

Db 365 SQLEFNKPYW 373

RESULT 26

SI5620

L1 protein - human papillomavirus type 2a

C:Species: human papillomavirus type 2a

A:Note: host Homo sapiens (man)

C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 17-Feb-1994

C:Accession: SI5620

R:Hirsch-Behnam, A.; Delius, H.; de Villiers, E.M.

Virus Res. 18, 81-98, 1990

A:Title: A comparative sequence analysis of two human papillomavirus (HPV) types 2a and

A:Reference number: SI5614; MUID:91188699; PMID:1964523

A:Accession: SI5620

A:Molecule type: DNA

A:Residues: 1-510 <HIR>

A:Cross-references: EMBL:X55964

C:Superfamily: papillomavirus L1 protein

C:Keywords: late protein

Query Match 89.1%; Score 49; DB 1; Length 510;
Best Local Similarity 87.5%; Pred. No. 0.2; 0; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 0;

QY 2 QIFNKPYW 9

```

Db          316 QLFNKPW 323
      |:|||||
RESULT 27
S15627
L1 protein - human papillomavirus type 57
C:Species: human papillomavirus type 57
A:Note: host Homo sapiens (man)
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
C:Accession: S15627
R:Hirsch-Behnam, A.; Delius, H.; de Villiers, E.M.
Virus Res. 18, 81-98, 1990
A:Title: A comparative sequence analysis of two human papillomavirus (HPV) types 2a and
A:Reference number: S15614; MUID:91188699; PMID:1964523
A:Accession: S15627
A:Molecule type: DNA
A:Residues: 1-510 <HNR>
A:Cross-references: EMBL:X55965; NID:G60882; PIDN:CAA39436.1; PID:G60889
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

      Query Match      89.1%; Score 49; DB 1; Length 510;
      Best Local Similarity 87.5%; Pred. No. 0.2;
      Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy          2 QIFNKPW 9
      |:|||||
Db          317 QLFNKPW 324

RESULT 28
S36554
L1 protein (alternative) - human papillomavirus type 3
C:Species: human papillomavirus type 3
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36554
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36554
A:Molecule type: DNA
A:Residues: 1-532 <DEL>
A:Cross-references: EMBL:X74462; NID:G397005; PIDN:CAA52474.1; PID:G397012
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

      Query Match      89.1%; Score 49; DB 2; Length 532;
      Best Local Similarity 87.5%; Pred. No. 0.21;
      Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy          2 QIFNKPW 9
      |:|||||
Db          335 QLFNKPW 342

RESULT 29
S36502
L1 protein (alternative) - human papillomavirus type 27
C:Species: human papillomavirus type 27
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36502
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36502
A:Molecule type: DNA
A:Residues: 1-594 <DEL>
A:Cross-references: EMBL:X74473; NID:G396964; PIDN:CAA52541.1; PID:G396971
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

```

```

      Query Match      89.1%; Score 49; DB 2; Length 594;
      Best Local Similarity 87.5%; Pred. No. 0.24;
      Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy          2 QIFNKPW 9
      |:|||||
Db          410 QLFNKPW 417

RESULT 30
P1WLB
L1 protein - bovine papillomavirus type 1
C:Species: bovine papillomavirus type 1
C>Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 28-Jul-2000
C:Accession: A03644
R:Chen, E.Y.; Howley, P.M.; Levinson, A.D.; Seeburg, P.H.
Nature 299, 529-534, 1982
A:Title: The primary structure and genetic organization of the bovine papillomavirus type
A:Reference number: A93289; MUID:83012974; PMID:6289124
A:Accession: A03644
A:Molecule type: DNA
A:Residues: 1-495 <CHE>
A:Cross-references: GB:X02346; GB:J02044; GB:M24622; GB:X00473; NID:G60965; PIDN:CAB4651
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

```

```

      Query Match      87.3%; Score 48; DB 1; Length 495;
      Best Local Similarity 87.5%; Pred. No. 0.31;
      Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy          2 QIFNKPW 9
      |:|||||
Db          302 QIFNRPW 309

RESULT 31
P1WLB2
L1 protein - bovine papillomavirus type 2
C:Species: bovine papillomavirus type 2
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 12-Jun-1998
C:Accession: A31169
R:Groff, D.E.; Mitra, R.; Lancaster, W.D.
submitted to GenBank, May 1988
A:Reference number: A94519
A:Accession: A31169
A:Molecule type: DNA
A:Residues: 1-501 <GRO>
A:Cross-references: GB:X01768; GB:M24326; NID:G60859
C:Superfamily: papillomavirus L1 protein
C:Keywords: glycoprotein; late protein
F:214,342,489/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

```

      Query Match      87.3%; Score 48; DB 1; Length 501;
      Best Local Similarity 87.5%; Pred. No. 0.31;
      Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy          2 QIFNKPW 9
      |:|||||
Db          306 QIFNRPW 313

RESULT 32
S48785
L1 protein - human papillomavirus (isolate vs93-1) (fragment)
N:Alternate names: capsid protein L1
C:Species: human papillomavirus
A:Variety: isolate vs93-1
C>Date: 10-Dec-1994 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C:Accession: S48785
R:Shamanin, V.; Glover, M.; Rausch, C.; Proby, C.; Leigh, I.M.; zur Hause, H.; Villiers,
submitted to the EMBL Data Library, June 1994
A:Description: Specific types of HPV found in benign proliferations and carcinomas of the

```

A;Reference number: S48779
A;Accession: S48785
A;Molecule type: DNA
A;Residues: 1-215 <SHA>
A;Cross-references: EMBL:X79947; NID:g562322; PIDN:CAA56295.1; PID:g562323
A;Experimental source: isolate vs93-1
C;Superfamily: papillomavirus L1 protein
C;Keywords: late protein

Query Match 85.5%; Score 47; DB 2; Length 215;
Best Local Similarity 88.9%; Pred. No. 0.19;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
|||||
DB 197 AQIFNKSYW 205

RESULT 33
S48781
L1 protein - human papillomavirus (isolate vs40-7) (fragment)
N;Alternate names: capsid protein L1
C;Species: human papillomavirus
A;Variety: isolate vs40-7
C;Date: 10-Dec-1994 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C;Accession: S48781
R;Shamanin, V.; Glover, M.; Rausch, C.; Proby, C.; Leigh, I.M.; zur Hause, H.; Villiers,
submitted to the EMBL Data Library, June 1994
A;Description: Specific types of HPV found in benign proliferations and carcinomas of th
A;Reference number: S48779
A;Accession: S48781
A;Molecule type: DNA
A;Residues: 1-220 <SHA>
A;Cross-references: EMBL:X79942; NID:g562314; PIDN:CAA56290.1; PID:g562315
A;Experimental source: isolate vs40-7
C;Superfamily: papillomavirus L1 protein
C;Keywords: late protein

Query Match 83.6%; Score 46; DB 2; Length 220;
Best Local Similarity 66.7%; Pred. No. 0.31;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
|||||
DB 204 AOLFNRPFW 212

RESULT 34
S48782
L1 protein - human papillomavirus (isolate vs42-1) (fragment)
N;Alternate names: capsid protein L1
C;Species: human papillomavirus
A;Variety: isolate vs42-1
C;Date: 10-Dec-1994 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C;Accession: S48782
R;Shamanin, V.; Glover, M.; Rausch, C.; Proby, C.; Leigh, I.M.; zur Hause, H.; Villiers,
submitted to the EMBL Data Library, June 1994
A;Description: Specific types of HPV found in benign proliferations and carcinomas of th
A;Reference number: S48779
A;Accession: S48782
A;Molecule type: DNA
A;Residues: 1-221 <SHA>
A;Cross-references: EMBL:X79943; NID:g562316; PIDN:CAA56291.1; PID:g562317
A;Experimental source: isolate vs42-1
C;Superfamily: papillomavirus L1 protein
C;Keywords: late protein

Query Match 83.6%; Score 46; DB 2; Length 221;
Best Local Similarity 66.7%; Pred. No. 0.31;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
|||||

A;Reference number: S48779
A;Accession: S48785
A;Molecule type: DNA
A;Residues: 1-215 <SHA>
A;Cross-references: EMBL:X79947; NID:g562322; PIDN:CAA56295.1; PID:g562323
A;Experimental source: isolate vs93-1
C;Superfamily: papillomavirus L1 protein
C;Keywords: late protein

Query Match 85.5%; Score 47; DB 2; Length 215;
Best Local Similarity 88.9%; Pred. No. 0.19;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
|||||
DB 197 AQIFNKSYW 205

RESULT 33
S48781
L1 protein - human papillomavirus (isolate vs40-7) (fragment)
N;Alternate names: capsid protein L1
C;Species: human papillomavirus
A;Variety: isolate vs40-7
C;Date: 10-Dec-1994 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C;Accession: S48781
R;Shamanin, V.; Glover, M.; Rausch, C.; Proby, C.; Leigh, I.M.; zur Hause, H.; Villiers,
submitted to the EMBL Data Library, June 1994
A;Description: Specific types of HPV found in benign proliferations and carcinomas of th
A;Reference number: S48779
A;Accession: S48781
A;Molecule type: DNA
A;Residues: 1-220 <SHA>
A;Cross-references: EMBL:X79942; NID:g562314; PIDN:CAA56290.1; PID:g562315
A;Experimental source: isolate vs40-7
C;Superfamily: papillomavirus L1 protein
C;Keywords: late protein

Query Match 83.6%; Score 46; DB 2; Length 220;
Best Local Similarity 66.7%; Pred. No. 0.31;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
|||||
DB 204 AOLFNRPFW 212

RESULT 34
S48782
L1 protein - human papillomavirus (isolate vs42-1) (fragment)
N;Alternate names: capsid protein L1
C;Species: human papillomavirus
A;Variety: isolate vs42-1
C;Date: 10-Dec-1994 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C;Accession: S48782
R;Shamanin, V.; Glover, M.; Rausch, C.; Proby, C.; Leigh, I.M.; zur Hause, H.; Villiers,
submitted to the EMBL Data Library, June 1994
A;Description: Specific types of HPV found in benign proliferations and carcinomas of th
A;Reference number: S48779
A;Accession: S48782
A;Molecule type: DNA
A;Residues: 1-221 <SHA>
A;Cross-references: EMBL:X79943; NID:g562316; PIDN:CAA56291.1; PID:g562317
A;Experimental source: isolate vs42-1
C;Superfamily: papillomavirus L1 protein
C;Keywords: late protein

Query Match 83.6%; Score 46; DB 2; Length 221;
Best Local Similarity 66.7%; Pred. No. 0.31;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
|||||

DB 203 AOLFNRPFW 211

RESULT 35
S48786
L1 protein - human papillomavirus (isolate cr148-59) (fragment)
N;Alternate names: capsid protein L1
C;Species: human papillomavirus
C;Date: 10-Dec-1994 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C;Accession: S48786
R;Shamanin, V.; Glover, M.; Rausch, C.; Proby, C.; Leigh, I.M.; zur Hause, H.; Villiers,
submitted to the EMBL Data Library, June 1994
A;Description: Specific types of HPV found in benign proliferations and carcinomas of th
A;Reference number: S48779
A;Accession: S48786
A;Molecule type: DNA
A;Residues: 1-222 <SHA>
A;Cross-references: EMBL:X79948; NID:g562308; PIDN:CAA56296.1; PID:g562309
A;Experimental source: isolate cr148-59
C;Superfamily: papillomavirus L1 protein
C;Keywords: late protein

Query Match 83.6%; Score 46; DB 2; Length 222;
Best Local Similarity 66.7%; Pred. No. 0.31;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
|||||
DB 204 AOLFNRPFW 212

RESULT 36
S48779
L1 protein - human papillomavirus (isolate vs102-4) (fragment)
N;Alternate names: capsid protein L1
C;Species: human papillomavirus
A;Variety: isolate vs102-4
C;Date: 10-Dec-1994 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C;Accession: S48779
R;Shamanin, V.; Glover, M.; Rausch, C.; Proby, C.; Leigh, I.M.; zur Hause, H.; Villiers,
submitted to the EMBL Data Library, June 1994
A;Description: Specific types of HPV found in benign proliferations and carcinomas of th
A;Reference number: S48779
A;Accession: S48779
A;Molecule type: DNA
A;Residues: 1-224 <SHA>
A;Cross-references: EMBL:X79946; NID:g562310; PIDN:CAA56294.1; PID:g562311
A;Experimental source: isolate vs102-4
C;Superfamily: papillomavirus L1 protein
C;Keywords: late protein

Query Match 83.6%; Score 46; DB 2; Length 224;
Best Local Similarity 66.7%; Pred. No. 0.31;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
|||||
DB 206 AOLFNRPFW 214

RESULT 37
S48784
L1 protein - human papillomavirus (isolate vs92-1) (fragment)
N;Alternate names: capsid protein L1
C;Species: human papillomavirus
A;Variety: isolate vs92-1
C;Date: 10-Dec-1994 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C;Accession: S48784
R;Shamanin, V.; Glover, M.; Rausch, C.; Proby, C.; Leigh, I.M.; zur Hause, H.; Villiers,
submitted to the EMBL Data Library, June 1994
A;Description: Specific types of HPV found in benign proliferations and carcinomas of th
A;Reference number: S48779
A;Accession: S48784

```
A:Molecule type: DNA
A:Residues: 1-224 <SHA>
A:Cross-references: EMBL:X79949; NID:g562320; PIDN:CAA56297.1; PID:g562321
A:Experimental source: isolate vs92-1
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match      83.6%; Score 46; DB 2; Length 224;
Best Local Similarity 66.7%; Pred. No. 0.31; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 3;

Qy      1 AQIFNKPYW 9
Db      206 AQLFNRPFW 214

RESULT 38
S48780
L1 protein - human papillomavirus (isolate vs20-4) (fragment)
N:Alternate names: capsid protein L1
C:Species: human papillomavirus
A:Variety: isolate vs20-4
C:Date: 10-Dec-1994 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
R:Shamanin, V.; Glover, M.; Rausch, C.; Proby, C.; Leigh, I.M.; zur Hause, H.; Villiers,
submitted to the EMBL Data Library, June 1994
A:Description: Specific types of HPV found in benign proliferations and carcinomas of th
A:Reference number: S48779
A:Accession: S48780
A:Molecule type: DNA
A:Residues: 1-225 <SHA>
A:Cross-references: EMBL:X79941; NID:g562312; PIDN:CAA56289.1; PID:g562313
A:Experimental source: isolate vs20-4
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match      83.6%; Score 46; DB 2; Length 225;
Best Local Similarity 66.7%; Pred. No. 0.32; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 3;

Qy      1 AQIFNKPYW 9
Db      207 AQLFNRPFW 215

RESULT 39
S48783
L1 protein - human papillomavirus (isolate vs75-3) (fragment)
N:Alternate names: capsid protein L1
C:Species: human papillomavirus
A:Variety: isolate vs75-3
C:Date: 10-Dec-1994 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
R:Shamanin, V.; Glover, M.; Rausch, C.; Proby, C.; Leigh, I.M.; zur Hause, H.; Villiers,
submitted to the EMBL Data Library, June 1994
A:Description: Specific types of HPV found in benign proliferations and carcinomas of th
A:Reference number: S48779
A:Accession: S48783
A:Molecule type: DNA
A:Residues: 1-228 <SHA>
A:Cross-references: EMBL:X79945; NID:g562318; PIDN:CAA56293.1; PID:g562319
A:Experimental source: isolate vs75-3
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match      83.6%; Score 46; DB 2; Length 228;
Best Local Similarity 66.7%; Pred. No. 0.32; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 3;

Qy      1 AQIFNKPYW 9
Db      210 AQLFNRPFW 218
```

```
RESULT 40
PIWLEP
L1 protein - European elk papillomavirus
C:Species: European elk papillomavirus
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 11-May-2000
C:Accession: A94457; A94506; G29499
R:Eriksson, A.
unpublished results 1987, cited by GenBank
A:Reference number: A94457
A:Accession: A94457
A:Molecule type: DNA
A:Residues: 1-501 <ERI>
A:Cross-references: GB:M15953; NID:g333025; PIDN:AAA66861.1; PID:g484027
R:Pettersson, U.
submitted to GenBank, August 1987
A:Reference number: A94506
A:Accession: A94506
A:Molecule type: DNA
A:Residues: 1-501 <PET>
A:Cross-references: GB:M15953; NID:g333025; PIDN:AAA66861.1; PID:g484027
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match      83.6%; Score 46; DB 1; Length 501;
Best Local Similarity 75.0%; Pred. No. 0.74; Mismatches 2; Indels 0; Gaps 0;
Matches 6; Conservative 2;

Qy      2 QIFNKPYW 9
Db      304 QLFNRPYW 311

Search completed: March 3, 2004, 10:25:24
Job time : 14.5 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 10:23:24 ; Search time 27 Seconds
(without alignments)
70.384 Million cell updates/sec

Title: US-09-980-064-1

Perfect score: 55

Sequence: 1 AQIFNKPYW 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	502	14 US-10-224-999A-3469	Sequence 3469, Ap
2	55	100.0	505	9 US-09-162-904A-3	Sequence 3, Appli
3	55	100.0	505	9 US-09-162-904A-4	Sequence 4, Appli
4	55	100.0	505	9 US-09-824-017-2	Sequence 2, Appli
5	55	100.0	505	10 US-09-986-118A-2	Sequence 2, Appli
6	55	100.0	505	15 US-10-367-095-6	Sequence 6, Appli
7	55	100.0	531	10 US-09-991-053-1	Sequence 1, Appli
8	55	100.0	531	14 US-10-301-260A-1	Sequence 1, Appli
9	50	90.9	132	13 US-10-056-359-29	Sequence 29, Appli
10	50	90.9	132	13 US-10-056-360-29	Sequence 29, Appli
11	50	90.9	139	13 US-10-056-359-2	Sequence 2, Appli
12	50	90.9	139	13 US-10-056-360-2	Sequence 2, Appli
13	50	90.9	144	13 US-10-056-359-5	Sequence 5, Appli
14	50	90.9	144	13 US-10-056-360-5	Sequence 5, Appli
15	50	90.9	568	10 US-09-991-053-3	Sequence 3, Appli

16	50	90.9	568	14 US-10-301-260A-3	Sequence 3, Appli
17	49	89.1	136	13 US-10-056-359-8	Sequence 8, Appli
18	49	89.1	136	13 US-10-056-360-8	Sequence 8, Appli
19	49	89.1	141	13 US-10-056-359-26	Sequence 26, Appli
20	49	89.1	141	13 US-10-056-360-26	Sequence 26, Appli
21	48	87.3	495	14 US-10-305-765-2	Sequence 2, Appli
22	48	87.3	495	14 US-10-305-765-4	Sequence 2, Appli
23	48	87.3	495	14 US-10-305-633-2	Sequence 4, Appli
24	48	87.3	495	14 US-10-305-633-4	Sequence 4, Appli
25	47	85.5	138	13 US-10-056-359-14	Sequence 14, Appli
26	47	85.5	138	13 US-10-056-359-23	Sequence 23, Appli
27	47	85.5	138	13 US-10-056-360-14	Sequence 14, Appli
28	47	85.5	138	13 US-10-056-360-23	Sequence 23, Appli
29	46	83.6	142	13 US-10-056-359-20	Sequence 20, Appli
30	46	83.6	142	13 US-10-056-360-20	Sequence 20, Appli
31	46	83.6	480	11 US-09-864-408A-8574	Sequence 8574, Ap
32	46	83.6	503	9 US-09-134-377-2	Sequence 2, Appli
33	46	83.6	503	9 US-09-822-662-2	Sequence 2, Appli
34	42	76.4	145	13 US-10-056-359-11	Sequence 11, Appli
35	42	76.4	145	13 US-10-056-360-11	Sequence 11, Appli
36	38	69.1	137	13 US-10-056-359-17	Sequence 17, Appli
37	38	69.1	137	13 US-10-056-360-17	Sequence 17, Appli
38	38	69.1	359	9 US-09-828-523A-14	Sequence 14, Appli
39	38	69.1	359	10 US-09-966-521-12	Sequence 12, Appli
40	38	69.1	359	14 US-10-429-094-12	Sequence 12, Appli
41	38	69.1	370	9 US-09-828-523A-74	Sequence 74, Appli
42	38	69.1	370	10 US-09-966-521-78	Sequence 78, Appli
43	38	69.1	370	14 US-10-429-094-78	Sequence 78, Appli
44	38	69.1	508	9 US-09-134-377-1	Sequence 1, Appli
45	38	69.1	508	9 US-09-822-662-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-10-224-999A-3469
; Sequence 3469, Application US/10224999A
; Publication No. US20030171318A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Composition and Method for Treating Viral Infection
; FILE REFERENCE: 5004.01
; CURRENT APPLICATION NUMBER: US/10/224,999A
; CURRENT FILING DATE: 2003-03-03
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 3484
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3469
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Human papillomavirus type 35
US-10-224-999A-3469

Query Match 100.0%; Score 55; DB 14; Length 502;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
| | | | | | | | | |
Db 302 AQIFNKPYW 310

RESULT 2

US-09-162-904A-3
; Sequence 3, Application US/09162904A
; Patent No. US20020168372A1
; GENERAL INFORMATION:
; APPLICANT: Durst, Matthias

```

; APPLICANT: Gissmann, Lutz
; TITLE OF INVENTION: A DNA SEQUENCE ENCODING A PAPILLOMAVIRUS L1 PROTEIN CAPABLE
; FILE REFERENCE: 8484-068-999
; CURRENT APPLICATION NUMBER: US/09/162,904A
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 08/884,168
; PRIOR FILING DATE: 1997-06-27
; PRIOR APPLICATION NUMBER: 08/641,570
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/902,528
; PRIOR FILING DATE: 1993-07-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid sequence of L1 gen
; OTHER INFORMATION: HPV16 clone
US-09-162-904A-3

Query Match 100.0%; Score 55; DB 9; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.11; 0; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 AQIFNKPYW 9
    |||||
Db 304 AQIFNKPYW 312

RESULT 3
US-09-162-904A-4
; Sequence 4, Application US/09162904A
; Patent No. US20020168372A1
; GENERAL INFORMATION:
; APPLICANT: Durst, Matthias
; APPLICANT: Gissmann, Lutz
; TITLE OF INVENTION: A DNA SEQUENCE ENCODING A PAPILLOMAVIRUS L1 PROTEIN CAPABLE
; FILE REFERENCE: 8484-068-999
; CURRENT APPLICATION NUMBER: US/09/162,904A
; CURRENT FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 08/884,168
; PRIOR FILING DATE: 1997-06-27
; PRIOR APPLICATION NUMBER: 08/641,570
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/902,528
; PRIOR FILING DATE: 1993-07-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid sequence of L1 gen
; OTHER INFORMATION: HPV16 clone
US-09-162-904A-4

Query Match 100.0%; Score 55; DB 9; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.11; 0; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 AQIFNKPYW 9
    |||||
Db 304 AQIFNKPYW 312

RESULT 4
US-09-824-017-2
; Sequence 2, Application US/09824017
```

```

; Publication No. US20020197668A1
; GENERAL INFORMATION:
; APPLICANT: BURGER, Alexander
; APPLICANT: HALLEK, Michael
; TITLE OF INVENTION: PAPILLOMA VIRUS CAPSOMERE VACCINE
; FORMULATIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/824,017
; FILING DATE: 03-Apr-2001
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/026,896
; FILING DATE: 1998-02-20
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercock, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 37067/102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 505 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-824-017-2

Query Match 100.0%; Score 55; DB 9; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.11; 0; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 AQIFNKPYW 9
    |||||
Db 304 AQIFNKPYW 312

RESULT 5
US-09-986-118A-2
; Sequence 2, Application US/09986118A
; Publication No. US20030021806A1
; GENERAL INFORMATION:
; APPLICANT: BURGER, Alexander
; APPLICANT: HALLEK, Michael
; TITLE OF INVENTION: PAPILLOMA VIRUS CAPSOMERE VACCINE
; FORMULATIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/986,118A
; FILING DATE: 07-No. US20030021806A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/026,896
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercock, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 37067/102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 505 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-986-118A-2

Query Match      100.0%; Score 55; DB 10; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AQIFNKPW 9
Db      304 AQIFNKPW 312

RESULT 6
US-10-367-095-6
; Sequence 6, Application US/10367095
; Publication No. US2003022869A1
; GENERAL INFORMATION:
; APPLICANT: Robin A. Robinson
; TITLE OF INVENTION: No. US2003022869A1el Insect Cell Line
; FILE REFERENCE: 44149-1US1
; CURRENT APPLICATION NUMBER: US/10/367,095
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 60/356,119
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,161
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,118
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,133
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,157
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,156
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,123
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,113
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,154
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,135
; PRIOR FILING DATE: 2002-02-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: HPV-16 L1 protein
US-10-367-095-6
```

```
Query Match      100.0%; Score 55; DB 15; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AQIFNKPW 9
Db      304 AQIFNKPW 312

RESULT 7
US-09-991-053-1
; Sequence 1, Application US/0991053
; Publication No. US2003000332A1
; GENERAL INFORMATION:
; APPLICANT: Harrison, Stephen
; APPLICANT: Chen, Xiaojiang
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
; FILE REFERENCE: Harvard/Harrison 12687/1120
; CURRENT APPLICATION NUMBER: US/09/991,053
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-18
; PRIOR APPLICATION NUMBER: 60/125208
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/148544
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-09-991-053-1

Query Match      100.0%; Score 55; DB 10; Length 531;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AQIFNKPW 9
Db      330 AQIFNKPW 338

RESULT 8
US-10-301-260A-1
; Sequence 1, Application US/10301260A
; Publication No. US20030118609A1
; GENERAL INFORMATION:
; APPLICANT: Harrison, Stephen
; APPLICANT: Chen, Xiaojiang
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
; FILE REFERENCE: Harvard/Harrison 12687/1123
; CURRENT APPLICATION NUMBER: US/10/301,260A
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: 09/520,822
; PRIOR FILING DATE: 2000-03-18
; PRIOR APPLICATION NUMBER: 60/125208
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/148544
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-301-260A-1

Query Match      100.0%; Score 55; DB 14; Length 531;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 AQIFNKPYW 9
Db      112 SOLFNKPYW 120

RESULT 11
US-10-056-359-2
; Sequence 2, Application US/10056359
; Publication No. US2002011086SAI
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: Zur Hausen, Ethel-Michele
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; FILE REFERENCE: 8484-0037-999
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2001-05-08
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-359-2

Query Match      90.9%; Score 50; DB 13; Length 139;
Best Local Similarity 77.8%; Pred. No. 0.26;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 AQIFNKPYW 9
Db      119 AQLFNRPYW 127

RESULT 12
US-10-056-360-2
; Sequence 2, Application US/10056360
; Publication No. US2002011086SAI
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: Zur Hausen, Ethel-Michele
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; FILE REFERENCE: 8484-0037-999
; CURRENT FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-360-2

Query Match      90.9%; Score 50; DB 13; Length 139;
Best Local Similarity 77.8%; Pred. No. 0.26;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 AQIFNKPYW 9
Db      112 SOLFNKPYW 120

RESULT 9
US-10-056-359-29
; Sequence 29, Application US/10056359
; Publication No. US2002011086SAI
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: Zur Hausen, Ethel-Michele
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; FILE REFERENCE: 8484-0037-999
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/10/056,359
; PRIOR FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-359-29

Query Match      90.9%; Score 50; DB 13; Length 132;
Best Local Similarity 77.8%; Pred. No. 0.25;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 AQIFNKPYW 9
Db      112 SOLFNKPYW 120

RESULT 10
US-10-056-360-29
; Sequence 29, Application US/10056360
; Publication No. US2002011086SAI
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: Zur Hausen, Ethel-Michele
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; FILE REFERENCE: 8484-0037-999
; CURRENT FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/10/056,360
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-360-29

Query Match      90.9%; Score 50; DB 13; Length 132;
Best Local Similarity 77.8%; Pred. No. 0.25;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```
Db 119 AQLFNRPYW 127
||:|:|:|
||:|:|:|

RESULT 13
US-10-056-359-5
; Sequence 5, Application US/10056359
; Publication No. US20020110865A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,359
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-359-5

Query Match 90.9%; Score 50; DB 13; Length 144;
Best Local Similarity 77.8%; Pred. No. 0.27;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
||:|:|:|
||:|:|:|

Db 124 AQLFNRPYW 132
||:|:|:|
||:|:|:|

RESULT 14
US-10-056-360-5
; Sequence 5, Application US/10056360
; Publication No. US20020110865A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,360
; CURRENT FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-360-5

Query Match 90.9%; Score 50; DB 13; Length 144;
Best Local Similarity 77.8%; Pred. No. 0.27;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
||:|:|:|
||:|:|:|

Db 124 AQLFNRPYW 132
||:|:|:|
||:|:|:|

RESULT 15
US-09-991-053-3
; Sequence 3, Application US/0991053
; Publication No. US2003000332A1
; GENERAL INFORMATION:
; APPLICANT: Harrison, Stephen
; APPLICANT: Chen, Xiaojiang
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
; TITLE OF INVENTION: virus
; FILE REFERENCE: Harvard/Harrison 12687/1120
; CURRENT APPLICATION NUMBER: US/09/991,053
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-18
; PRIOR APPLICATION NUMBER: 60/125208
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/148544
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Human papillomavirus type 18
US-09-991-053-3

Query Match 90.9%; Score 50; DB 10; Length 568;
Best Local Similarity 77.8%; Pred. No. 0.95;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
||:|:|:|
||:|:|:|

Db 365 SOLFNKPYW 373

RESULT 16
US-10-301-260A-3
; Sequence 3, Application US/10301260A
; Publication No. US20030118609A1
; GENERAL INFORMATION:
; APPLICANT: Harrison, Stephen
; APPLICANT: Chen, Xiaojiang
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
; TITLE OF INVENTION: virus
; FILE REFERENCE: Harvard/Harrison 12687/1123
; CURRENT APPLICATION NUMBER: US/10/301,260A
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: 09/520,822
; PRIOR FILING DATE: 2000-03-18
; PRIOR APPLICATION NUMBER: 60/125208
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/148544
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Human papillomavirus type 18
US-10-301-260A-3

Query Match 90.9%; Score 50; DB 14; Length 568;
Best Local Similarity 77.8%; Pred. No. 0.95;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
||:|:|:|
||:|:|:|

Db 365 SOLFNKPYW 373
```

```
RESULT 17
US-10-056-359-8
; Sequence 8, Application US/10056359
; Publication No. US20020110865A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,359
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2001-05-08
; PRIOR FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1995-07-19
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-359-8
```

```
Query Match 89.1%; Score 49; DB 13; Length 136;
Best Local Similarity 87.5%; Pred. No. 0.38;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 QIFNKPW 9
Db 117 QLFNRPW 124
```

```
RESULT 18
US-10-056-360-8
; Sequence 8, Application US/10056360
; Publication No. US20020110866A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,360
; CURRENT FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1995-07-19
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-360-8
```

```
Query Match 89.1%; Score 49; DB 13; Length 136;
Best Local Similarity 87.5%; Pred. No. 0.38;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 QIFNKPW 9
Db 117 QLFNRPW 124
```

```
RESULT 19
```

```
US-10-056-359-26
; Sequence 26, Application US/10056359
; Publication No. US20020110865A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,359
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2001-05-08
; PRIOR FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1995-07-19
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-359-26
```

```
Query Match 89.1%; Score 49; DB 13; Length 141;
Best Local Similarity 77.8%; Pred. No. 0.39;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AQIFNKPW 9
Db 121 SQIFNRPW 129
```

```
RESULT 20
US-10-056-360-26
; Sequence 26, Application US/10056360
; Publication No. US20020110866A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,360
; CURRENT FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1995-07-19
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-360-26
```

```
Query Match 89.1%; Score 49; DB 13; Length 141;
Best Local Similarity 77.8%; Pred. No. 0.39;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AQIFNKPW 9
Db 121 SQIFNRPW 129
```

```
RESULT 21
US-10-305-765-2
```

```

; Sequence 2, Application US/10305765
; Publication No. US20030175907A1
; GENERAL INFORMATION:
; APPLICANT: FRAZER, Ian Hector
; APPLICANT: ZHOU, Jian
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY
; TITLE OF INVENTION: EXPRESSING A PROTEIN IN A TARGET CELL OR TISSUE
; FILE REFERENCE: 210338.0001/IUS
; CURRENT APPLICATION NUMBER: US/10/305,765
; CURRENT FILING DATE: 2002-11-27
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US/09/479,645
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: PCT/AU98/00530
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: AU P07765
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: AU P09467
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Bovine papillomavirus type 1
; FEATURE:
; OTHER INFORMATION: L1 open reading frame (wild-type)
US-10-305-765-2

Query Match      87.3%; Score 48; DB 14; Length 495;
Best Local Similarity 87.5%; Pred. No. 1.9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy      2 QIFNKPW 9
Db      302 QIFNRPW 309

RESULT 22
US-10-305-765-4
; Sequence 4, Application US/10305765
; Publication No. US20030175907A1
; GENERAL INFORMATION:
; APPLICANT: FRAZER, Ian Hector
; APPLICANT: ZHOU, Jian
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY
; TITLE OF INVENTION: EXPRESSING A PROTEIN IN A TARGET CELL OR TISSUE
; FILE REFERENCE: 210338.0001/IUS
; CURRENT APPLICATION NUMBER: US/10/305,765
; CURRENT FILING DATE: 2002-11-27
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US/09/479,645
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: PCT/AU98/00530
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: AU P07765
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: AU P09467
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Bovine
; OTHER INFORMATION: papillomavirus type 1 L1 open reading frame
; OTHER INFORMATION: (humanized)
; FEATURE:
; OTHER INFORMATION: Wild-type codons replaced with synonymous codons
; OTHER INFORMATION: used at relatively high frequency by human genes
US-10-305-765-4

Query Match      87.3%; Score 48; DB 14; Length 495;

```

```

Best Local Similarity 87.5%; Pred. No. 1.9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy      2 QIFNKPW 9
Db      302 QIFNRPW 309

RESULT 23
US-10-305-633-2
; Sequence 2, Application US/10305633
; Publication No. US20030182674A1
; GENERAL INFORMATION:
; APPLICANT: FRAZER, Ian Hector
; APPLICANT: ZHOU, Jian
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY
; TITLE OF INVENTION: EXPRESSING A PROTEIN IN A TARGET CELL OR TISSUE
; FILE REFERENCE: 210338.0001/IUS
; CURRENT APPLICATION NUMBER: US/10/305,633
; CURRENT FILING DATE: 2002-11-27
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US/09/479,645
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: PCT/AU98/00530
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: AU P07765
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: AU P09467
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Bovine papillomavirus type 1
; FEATURE:
; OTHER INFORMATION: L1 open reading frame (wild-type)
US-10-305-633-2

Query Match      87.3%; Score 48; DB 14; Length 495;
Best Local Similarity 87.5%; Pred. No. 1.9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy      2 QIFNKPW 9
Db      302 QIFNRPW 309

RESULT 24
US-10-305-633-4
; Sequence 4, Application US/10305633
; Publication No. US20030182674A1
; GENERAL INFORMATION:
; APPLICANT: FRAZER, Ian Hector
; APPLICANT: ZHOU, Jian
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY
; TITLE OF INVENTION: EXPRESSING A PROTEIN IN A TARGET CELL OR TISSUE
; FILE REFERENCE: 210338.0001/IUS
; CURRENT APPLICATION NUMBER: US/10/305,633
; CURRENT FILING DATE: 2002-11-27
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US/09/479,645
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: PCT/AU98/00530
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: AU P07765
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: AU P09467
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

```
;
; FEATURE: Description of Artificial Sequence: Bovine
; OTHER INFORMATION: papillomavirus type 1 L1 open reading frame
; OTHER INFORMATION: (humanized)
; FEATURE:
; OTHER INFORMATION: Wild-type codons replaced with synonymous codons
; OTHER INFORMATION: used at relatively high frequency by human genes
US-10-305-633-4

Query Match      87.3%; Score 48; DB 14; Length 495;
Best Local Similarity 87.5%; Pred. No. 1.9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QIFNKPYPW 9
   |||:||||
Db 302 QIFNRPYW 309

RESULT 25
US-10-056-359-14
; Sequence 14, Application US/10056359
; Publication No. US20020110865A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,359
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 138.
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-359-14

Query Match      85.5%; Score 47; DB 13; Length 138;
Best Local Similarity 66.7%; Pred. No. 0.86;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYPW 9
   |||:||||
Db 118 SOLFNRPYW 126

RESULT 26
US-10-056-359-23
; Sequence 23, Application US/10056359
; Publication No. US20020110865A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,359
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
```

```
;
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-359-23

Query Match      85.5%; Score 47; DB 13; Length 138;
Best Local Similarity 66.7%; Pred. No. 0.86;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AOIFNKPYPW 9
   |||:||||
Db 118 SOLFNRPYW 126

RESULT 27
US-10-056-360-14
; Sequence 14, Application US/10056360
; Publication No. US20020110866A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: Zur Hausen, Harald
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,360
; CURRENT FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-360-14

Query Match      85.5%; Score 47; DB 13; Length 138;
Best Local Similarity 66.7%; Pred. No. 0.86;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AOIFNKPYPW 9
   |||:||||
Db 118 SOLFNRPYW 126

RESULT 28
US-10-056-360-23
; Sequence 23, Application US/10056360
; Publication No. US20020110866A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: Zur Hausen, Harald
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,360
; CURRENT FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
```



```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-360-23

Query Match      85.5%; Score 47; DB 13; Length 138;
Best Local Similarity 66.7%; Pred. No. 0.86;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AQIFNKPYW 9
      :|:|:|:|
Db      118 QLFNRPYW 126

RESULT 29
US-10-056-359-20
; Sequence 20, Application US/10056359
; Publication No. US20020110866A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,359
; CURRENT FILING DATE: 2002-01-23
; PRIOR FILING DATE: US/09/628,099
; PRIOR FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-359-20

Query Match      83.6%; Score 46; DB 13; Length 142;
Best Local Similarity 75.0%; Pred. No. 1.3;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 QIFNKPYW 9
      :|:|:|:|
Db      123 QLFNRPYW 130

RESULT 30
US-10-056-360-20
; Sequence 20, Application US/10056360
; Publication No. US20020110866A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,360
; CURRENT FILING DATE: 2000-01-23
; PRIOR FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 20
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-360-20

Query Match      83.6%; Score 46; DB 13; Length 142;
Best Local Similarity 75.0%; Pred. No. 1.3;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 QIFNKPYW 9
      :|:|:|:|
Db      123 QLFNRPYW 130

RESULT 31
US-09-864-408A-8574
; Sequence 8574, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8574
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-8574

Query Match      83.6%; Score 46; DB 11; Length 480;
Best Local Similarity 66.7%; Pred. No. 4.1;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AQIFNKPYW 9
      :|:|:|:|
Db      280 AQLFNRPFW 288

RESULT 32
US-09-134-377-2
; Sequence 2, Application US/09134377
; Patent No. US20020015710A1
; GENERAL INFORMATION:
; APPLICANT: SCHLEGEL, C. Richard
; APPLICANT: JENSON, A. Bennett
; APPLICANT: GHIM, Shin-Je
; TITLE OF INVENTION: PAPILLOMAVIRUS VACCINES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/134,377
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/724,281
```

```

Best Local Similarity 75.0%; Pred No. 4.3;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 QIFNKPW 9
Db      307 QLFNRPW 314

RESULT 34
US-10-056-359-11
; Sequence 11, Application US/10056359
; Publication No. US20020110865A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,359
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-359-11

Query Match 76.4%; Score 42; DB 13; Length 145;
Best Local Similarity 62.5%; Pred. No. 6.8;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 QIFNKPW 9
Db      126 QLFNRPW 133

RESULT 35
US-10-056-360-11
; Sequence 11, Application US/10056360
; Publication No. US20020110866A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,360
; CURRENT FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-360-11

Query Match 76.4%; Score 42; DB 13; Length 145;
Best Local Similarity 62.5%; Pred. No. 6.8;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QIFNKPW 9
|:|:|:|:
Db 126 QLFNRPWF 133

RESULT 36

US-10-056-359-17
; Sequence 17, Application US/10056359
; Publication No. US20020110865A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056.359
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-359-17

Query Match 69.1%; Score 38; DB 13; Length 137;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QIFNKPW 9
|:|:|:|:
Db 118 QIFNRPYF 125

RESULT 37

US-10-056-360-17
; Sequence 17, Application US/10056360
; Publication No. US20020110866A1
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056.360
; CURRENT FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-360-17

Query Match 69.1%; Score 38; DB 13; Length 137;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QIFNKPW 9
|:|:|:|:
Db 118 QIFNRPYF 125

RESULT 38

US-09-828-523A-14
; Sequence 14, Application US/09828523A
; Patent No. US20020168697A1
; GENERAL INFORMATION:
; APPLICANT: The Pharmacia & Upjohn Company
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 268.62120101
; CURRENT APPLICATION NUMBER: US/09/828,523A
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/266,327
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-828-523A-14

Query Match 69.1%; Score 38; DB 9; Length 359;
Best Local Similarity 75.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QIFNKPW 9
|:|:|:|:
Db 111 QSFNLPW 118

RESULT 39

US-09-966-521-12
; Sequence 12, Application US/09966521
; Publication No. US20030087321A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 6212.N2
; CURRENT APPLICATION NUMBER: US/09/966,521
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-966-521-12

Query Match 69.1%; Score 38; DB 10; Length 359;
Best Local Similarity 75.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QIFNKPW 9
|:|:|:|:
Db 111 QSFNLPW 118

RESULT 40

US-10-429-094-12
; Sequence 12, Application US/10429094
; Publication No. US20030180821A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl

; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 00774 US1 CN1
; CURRENT APPLICATION NUMBER: US/10/429,094
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 09/966,521
; PRIOR FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-429-094-12

Query Match 69.1%; Score 38; DB 14; Length 359;
Best Local Similarity 75.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 QIFNKPYW 9
| | | | |
Db 111 QSFNLPYW 118

Search completed: March 3, 2004, 10:27:16
Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 10:22:54 ; Search time 16.5 Seconds
(without alignments)
28.160 Million cell updates/sec

Title: US-09-980-064-1
Perfect score: 55
Sequence: 1 AQIFNKPW 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	55	100.0	497	4 US-09-413-611A-6
2	55	100.0	499	4 US-09-413-611A-8
3	55	100.0	505	3 US-08-944-368A-2
4	55	100.0	505	4 US-09-820-764-2
5	55	100.0	505	4 US-09-986-118A-2
6	55	100.0	505	4 US-09-824-017-2
7	55	100.0	531	4 US-09-520-822A-1
8	53	96.4	128	4 US-09-319-056B-14
9	53	96.4	128	4 US-09-319-056B-17
10	53	96.4	500	2 US-08-987-519-1
11	53	96.4	500	2 US-08-987-519-2
12	53	96.4	501	2 US-08-987-519-3
13	53	96.4	532	3 US-08-737-336-6
14	51	92.7	15	4 US-09-947-372A-34
15	50	90.9	132	4 US-09-000-266-29
16	50	90.9	132	4 US-09-628-099-29
17	50	90.9	132	4 US-10-056-360-29
18	50	90.9	132	4 US-10-056-359-29
19	50	90.9	139	4 US-09-000-266-2
20	50	90.9	139	4 US-09-628-099-2
21	50	90.9	139	4 US-10-056-360-2
22	50	90.9	139	4 US-10-056-359-2
23	50	90.9	144	4 US-09-000-266-5
24	50	90.9	144	4 US-09-628-099-5
25	50	90.9	144	4 US-10-056-360-5
26	50	90.9	144	4 US-10-056-359-5
27	50	90.9	507	2 US-08-409-122-2

28	50	90.9	507	2	US-08-408-669-2	Sequence 2, Appli
29	50	90.9	568	4	US-09-520-822A-3	Sequence 3, Appli
30	49	89.1	127	4	US-09-319-056B-11	Sequence 11, Appli
31	49	89.1	136	4	US-09-000-266-8	Sequence 8, Appli
32	49	89.1	136	4	US-09-628-099-8	Sequence 8, Appli
33	49	89.1	136	4	US-10-056-360-8	Sequence 8, Appli
34	49	89.1	136	4	US-10-056-359-8	Sequence 8, Appli
35	49	89.1	141	4	US-09-000-266-26	Sequence 26, Appli
36	49	89.1	141	4	US-09-628-099-26	Sequence 26, Appli
37	49	89.1	141	4	US-10-056-360-26	Sequence 26, Appli
38	49	89.1	141	4	US-10-056-359-26	Sequence 26, Appli
39	48	87.3	495	4	US-09-479-645A-2	Sequence 2, Appli
40	48	87.3	495	4	US-09-479-645A-4	Sequence 4, Appli
41	48	87.3	964	4	US-08-484-791-2	Sequence 2, Appli
42	47	85.5	136	4	US-09-319-056B-8	Sequence 8, Appli
43	47	85.5	138	4	US-09-000-266-14	Sequence 14, Appli
44	47	85.5	138	4	US-09-000-266-23	Sequence 23, Appli
45	47	85.5	138	4	US-09-628-099-14	Sequence 14, Appli

ALIGNMENTS

RESULT 1
US-09-413-611A-6
; Sequence 6, Application US/09413611A
; Patent No. 6380364
; GENERAL INFORMATION:
; APPLICANT: Mueller, Martin
; APPLICANT: Kast, Wjbe
; APPLICANT: Nieland, John
; APPLICANT: Velders, Markwin
; TITLE OF INVENTION: Chimeric Biotin-Binding Papillomavirus Protein
; FILE REFERENCE: 202325
; CURRENT APPLICATION NUMBER: US/09/413,611A
; CURRENT FILING DATE: 1999-10-06
; PRIOR FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:biotin-binding
; OTHER INFORMATION: papillomavirus protein 168
US-09-413-611A-6

Query Match	100.0%;	Score 55;	DB 4;	Length 497;
Best Local Similarity	100.0%;	Pred. NO. 0.038;		
Matches	9;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;
Qy	1	AQIFNKPW	9	
Db	304	AQIFNKPW	312	

RESULT 2
US-09-413-611A-8
; Sequence 8, Application US/09413611A
; Patent No. 6380364
; GENERAL INFORMATION:
; APPLICANT: Mueller, Martin
; APPLICANT: Kast, Wjbe
; APPLICANT: Nieland, John
; APPLICANT: Velders, Markwin
; TITLE OF INVENTION: Chimeric Biotin-Binding Papillomavirus Protein
; FILE REFERENCE: 202325
; CURRENT APPLICATION NUMBER: US/09/413,611A
; CURRENT FILING DATE: 1999-10-06
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 60/109,510
; PRIOR FILING DATE: 1998-11-23

```
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:biotin-binding
; OTHER INFORMATION: papillomavirus protein 169
US-09-413-611A-8

Query Match          100.0%; Score 55; DB 4; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.038; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY      1 AQIFNKPYW 9
      |||||
Db      304 AQIFNKPYW 312

RESULT 3
US-08-944-368A-2
; Sequence 2, Application US/08944368A
; Patent No. 6228368
; GENERAL INFORMATION:
; APPLICANT: Gissman, et al.
; TITLE OF INVENTION: Papilloma Virus Capsomere Vaccine
; TITLE OF INVENTION: Formulations and Methods of Use
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESS: Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,368A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27013/34028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 505 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-944-368A-2

Query Match          100.0%; Score 55; DB 3; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AQIFNKPYW 9
      |||||
Db      304 AQIFNKPYW 312

RESULT 4
US-09-820-764-2
; Sequence 2, Application US/09820764
```

```
; Patent No. 6352696
; GENERAL INFORMATION:
; APPLICANT: BURGER, Alexander
; HALLEK, Michael
; TITLE OF INVENTION: PAPILLOMA VIRUS CAPSOMERE VACCINE
; FORMULATIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/820,764
; FILING DATE: 30-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/026,896
; FILING DATE: 20-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercock, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 37067/102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 505 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-820-764-2

Query Match          100.0%; Score 55; DB 4; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AQIFNKPYW 9
      |||||
Db      304 AQIFNKPYW 312

RESULT 5
US-09-986-118A-2
; Sequence 2, Application US/09986118A
; Patent No. 6562351
; GENERAL INFORMATION:
; APPLICANT: BURGER, Alexander
; HALLEK, Michael
; TITLE OF INVENTION: PAPILLOMA VIRUS CAPSOMERE VACCINE
; FORMULATIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/986,118A
;; FILING DATE: 07-No. 6562351-2001
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 09/026,896
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Sandercock, Colin G.
;; REGISTRATION NUMBER: 31,298
;; REFERENCE/DOCKET NUMBER: 37067/102
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 672-5300
;; TELEFAX: (202) 672-5399
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 505 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-986-118A-2

Query Match 100.0%; Score 55; DB 4; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPW 9
Db 304 AQIFNKPW 312
|||||

RESULT 6
US-09-824-017-2
; Sequence 2, Application US/09824017
; Patent No. 6649167
; GENERAL INFORMATION:
; APPLICANT: BURGER, Alexander
; TITLE OF INVENTION: PAPILLOMA VIRUS CAPSOMERE VACCINE
; FORMULATIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/824,017
; FILING DATE: 03-Apr-2001
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/026,896
; FILING DATE: 1998-02-20
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercock, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 37067/102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 505 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-824-017-2

Query Match 100.0%; Score 55; DB 4; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPW 9
Db 304 AQIFNKPW 312
|||||

RESULT 7
US-09-520-822A-1
; Sequence 1, Application US/09520822A
; Patent No. 6551597
; GENERAL INFORMATION:
; APPLICANT: Harrison, Stephen
; APPLICANT: Chen, Xiaojiang
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
; TITLE OF INVENTION: virus
; FILE REFERENCE: Harvard/Harrison 12687/1120
; CURRENT APPLICATION NUMBER: US/09/520,822A
; CURRENT FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/125208
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/148544
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-09-520-822A-1

Query Match 100.0%; Score 55; DB 4; Length 531;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPW 9
Db 330 AQIFNKPW 338
|||||

RESULT 8
US-09-319-056B-14
; Sequence 14, Application US/09319056B
; Patent No. 6413522
; GENERAL INFORMATION:
; APPLICANT: de Villiers-zur Hausen, Ethel-Michele
; APPLICANT: zur Hausen, Harald
; APPLICANT: Laverne, Donna
; APPLICANT: Benton, Claire
; TITLE OF INVENTION: PAPILLOMA VIRUSES, PRODUCTS FOR THE
; TITLE OF INVENTION: DETECTION THEREOF AS WELL AS FOR TREATING DISEASES CAUSED BY
; TITLE OF INVENTION: THEM
; FILE REFERENCE: 8484-0086-999
; CURRENT APPLICATION NUMBER: US/09/319,056B
; CURRENT FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: PCT/DE97/02659
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: DE 196 48 962.8
; PRIOR FILING DATE: 1996-11-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Papilloma virus
US-09-319-056B-14

```
Query Match          96.4%; Score 53; DB 4; Length 128;
Best Local Similarity 88.9%; Pred. No. 0.022;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 AQIFNKPYW 9
Db      110 AQLFNKPYW 118

RESULT 9
US-09-319-056B-17
; Sequence 17, Application US/09319056B
; Patent No. 6413522
; GENERAL INFORMATION:
; APPLICANT: de Villiers-zur Hausen, Ethel-Michele
; APPLICANT: zur Hausen, Harald
; APPLICANT: Lavergne, Donna
; APPLICANT: Beniton, Claire
; TITLE OF INVENTION: PAPILLOMA VIRUSES, PRODUCTS FOR THE
; TITLE OF INVENTION: DETECTION THEREOF AS WELL AS FOR TREATING DISEASES CAUSED BY
; TITLE OF INVENTION: THEM
; FILE REFERENCE: 8484-0086-999
; CURRENT APPLICATION NUMBER: US/09/319,056B
; CURRENT FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: PCT/DE97/02659
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: DE 196 48 962.8
; PRIOR FILING DATE: 1996-11-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Papilloma virus
US-09-319-056B-17

Query Match          96.4%; Score 53; DB 4; Length 128;
Best Local Similarity 88.9%; Pred. No. 0.022;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 AQIFNKPYW 9
Db      110 AQLFNKPYW 118

RESULT 10
US-08-987-519-1
; Sequence 1, Application US/08987519
; Patent No. 5952216
; GENERAL INFORMATION:
; APPLICANT: Ludmerer, Steven
; TITLE OF INVENTION: Synthetic HPV16 Virus-Like Particles
; FILE REFERENCE: 19853
; CURRENT APPLICATION NUMBER: US/08/987,519
; CURRENT FILING DATE: 1997-12-09
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Human Papillomavirus
US-08-987-519-1

Query Match          96.4%; Score 53; DB 2; Length 500;
Best Local Similarity 88.9%; Pred. No. 0.087;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 AQIFNKPYW 9
Db      301 AQLFNKPYW 309

RESULT 11
```

```
US-08-987-519-2
; Sequence 2, Application US/08987519
; Patent No. 5952216
; GENERAL INFORMATION:
; APPLICANT: Ludmerer, Steven
; TITLE OF INVENTION: Synthetic HPV16 Virus-Like Particles
; FILE REFERENCE: 19853
; CURRENT APPLICATION NUMBER: US/08/987,519
; CURRENT FILING DATE: 1997-12-09
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Human Papillomavirus
US-08-987-519-2

Query Match          96.4%; Score 53; DB 2; Length 500;
Best Local Similarity 88.9%; Pred. No. 0.087;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 AQIFNKPYW 9
Db      300 AQLFNKPYW 308

RESULT 12
US-08-987-519-3
; Sequence 3, Application US/08987519
; Patent No. 5952216
; GENERAL INFORMATION:
; APPLICANT: Ludmerer, Steven
; TITLE OF INVENTION: Synthetic HPV16 Virus-Like Particles
; FILE REFERENCE: 19853
; CURRENT APPLICATION NUMBER: US/08/987,519
; CURRENT FILING DATE: 1997-12-09
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (0)...(0)
; OTHER INFORMATION: Consensus Sequence
; OTHER INFORMATION: At 28 X = K or T
; OTHER INFORMATION: At 49 X = Y or F
; OTHER INFORMATION: At 53 X = K or R
; OTHER INFORMATION: At 54 X = V or A
; OTHER INFORMATION: At 119 X = L or F
; OTHER INFORMATION: At 131 X = G or S
; OTHER INFORMATION: At 132 X = Y or K
; OTHER INFORMATION: At 170 X = T or K
; OTHER INFORMATION: At 173 X = S or T
; OTHER INFORMATION: At 176 X = S or P
; OTHER INFORMATION: At 179 X = N or A
; OTHER INFORMATION: At 219 X = L or I
; OTHER INFORMATION: At 225 X = V or T
; OTHER INFORMATION: At 246 X = Y or F
; OTHER INFORMATION: At 263 X = T or E
; OTHER INFORMATION: At 271 X = D or T
; OTHER INFORMATION: At 273 X = L or I
; OTHER INFORMATION: At 274 X = V or I
; OTHER INFORMATION: At 277 X = G or S
; OTHER INFORMATION: At 278 X = N or G
; OTHER INFORMATION: At 281 X = S or T
; OTHER INFORMATION: At 284 X = A or G
; OTHER INFORMATION: At 290 X = H or N
; OTHER INFORMATION: At 325 X = H or Q
; OTHER INFORMATION: At 346 X = S or T
; OTHER INFORMATION: At 347 X = K or T
; OTHER INFORMATION: At 348 X = A or S
```


OTHER INFORMATION: AT 366 X = F OR Y
OTHER INFORMATION: AT 434 X = Q OR P
OTHER INFORMATION: AT 439 X = D OR N
OTHER INFORMATION: AT 440 X = M OR L
OTHER INFORMATION: AT 458 X = F OR Y
OTHER INFORMATION: AT 474 X = T OR S
OTHER INFORMATION: AT 476 X = A OR I
OTHER INFORMATION: AT 480 X = I OR V
OTHER INFORMATION: AT 488 X = P OR A
OTHER INFORMATION: AT 490 X = T OR A
OTHER INFORMATION: AT 497 X = T OR A
OTHER INFORMATION: AT 501 X = K OR R
US-08-987-519-3

Query Match 96.4%; Score 53; DB 2; Length 501;
Best Local Similarity 88.9%; Pred. No. 0.088;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
Db 301 AQLFNKPYW 309

RESULT 13

US-08-737-336-6
Sequence 6, Application US/08737336
Patent No. 6013262
GENERAL INFORMATION:
APPLICANT: FRAZER, Ian
APPLICANT: ZHOU, Jian
TITLE OF INVENTION: RECOMBINANT PAPILLOMA VIRUS L1
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,336
FILING DATE: 16-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/AU95/00292
FILING DATE: 17-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PM 5667
FILING DATE: 17-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: SAXE, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 65064/118
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-737-336-6

Query Match 96.4%; Score 53; DB 3; Length 532;
Best Local Similarity 88.9%; Pred. No. 0.093;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
Db 332 AQLFNKPYW 340

RESULT 14

US-09-947-372A-34
Sequence 34, Application US/09947372A
Patent No. 6613557
GENERAL INFORMATION:
APPLICANT: FRAZER, Ian
APPLICANT: ZHOU, Jian
TITLE OF INVENTION: PAPILLOMAVIRUS VACCINE
FILE REFERENCE: 065064/0137
CURRENT APPLICATION NUMBER: US/09/947,372A
CURRENT FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 08/185,928
PRIOR FILING DATE: 1994-01-19
PRIOR APPLICATION NUMBER: PCT/AU92/02184
PRIOR FILING DATE: 1992-07-20
PRIOR APPLICATION NUMBER: AU PK7322
PRIOR FILING DATE: 1991-07-19
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 34
LENGTH: 15
TYPE: PRT
ORGANISM: Human papillomavirus type 16
US-09-947-372A-34

Query Match 92.7%; Score 51; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0056;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QIFNKPYW 9
Db 1 QIFNKPYW 8

RESULT 15

US-09-000-266-29
Sequence 29, Application US/09000266A
Patent No. 6322795
GENERAL INFORMATION:
APPLICANT: Shamanin, Vladimir
APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
APPLICANT: Zur Hausen, Harald
TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
FILE REFERENCE: 8484-0037-999
CURRENT APPLICATION NUMBER: US/09/000,266A
CURRENT FILING DATE: 1998-10-19
EARLIER APPLICATION NUMBER: PCT/DE96/01369
EARLIER FILING DATE: 1996-07-19
EARLIER APPLICATION NUMBER: DE P 195 26 386.3
EARLIER FILING DATE: 1995-07-19
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 29
LENGTH: 132
TYPE: PRT
ORGANISM: Papilloma virus
US-09-000-266-29

Query Match 90.9%; Score 50; DB 4; Length 132;
Best Local Similarity 77.8%; Pred. No. 0.078;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
Db 112 SOLFNKPYW 120

```
RESULT 16
US-09-628-099-29
; Sequence 29, Application US/09628099
; Patent No. 6368832
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Papilloma virus
US-09-628-099-29

Query Match          90.9%; Score 50; DB 4; Length 132;
Best Local Similarity 77.8%; Pred. No. 0.078;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
Db 112 SOLFNKPYW 120

RESULT 17
US-10-056-360-29
; Sequence 29, Application US/10056360
; Patent No. 6555345
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,360
; PRIOR FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-360-29

Query Match          90.9%; Score 50; DB 4; Length 132;
Best Local Similarity 77.8%; Pred. No. 0.078;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
Db 112 SOLFNKPYW 120
```

```
RESULT 19
US-09-000-266-2
; Sequence 2, Application US/09000266A
; Patent No. 6322795
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/09/000,266A
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: PCT/DE96/01369
; PRIOR FILING DATE: 1996-07-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Papilloma virus
US-09-000-266-2

Query Match          90.9%; Score 50; DB 4; Length 139;
Best Local Similarity 77.8%; Pred. No. 0.082;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
Db 119 AQLFNRPYW 127

RESULT 20
US-09-628-099-2
; Sequence 2, Application US/09628099
; Patent No. 6368832
; GENERAL INFORMATION:
```

```

; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Papilloma virus
US-09-628-099-2

Query Match          90.9%; Score 50; DB 4; Length 139;
Best Local Similarity 77.8%; Pred. No. 0.082;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AQIFNKPYW 9
Db      119 AQLFNRPYW 127

RESULT 21
US-10-056-360-2
; Sequence 2, Application US/10056360
; Patent No. 6555345
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,360
; CURRENT FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-360-2

Query Match          90.9%; Score 50; DB 4; Length 139;
Best Local Similarity 77.8%; Pred. No. 0.082;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AQIFNKPYW 9
Db      119 AQLFNRPYW 127

RESULT 22
US-10-056-359-2
; Sequence 2, Application US/10056359
; Patent No. 6562597
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,359
; CURRENT FILING DATE: 2000-01-23
; PRIOR APPLICATION NUMBER: US/09/628,099
; CURRENT FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-359-2

Query Match          90.9%; Score 50; DB 4; Length 139;
Best Local Similarity 77.8%; Pred. No. 0.082;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AQIFNKPYW 9
Db      119 AQLFNRPYW 127

RESULT 23
US-09-000-266-5
; Sequence 5, Application US/09000266A
; Patent No. 6322795
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/09/000,266A
; CURRENT FILING DATE: 1998-10-19
; EARLIER APPLICATION NUMBER: PCT/DE96/01369
; EARLIER FILING DATE: 1996-07-19
; EARLIER APPLICATION NUMBER: DE P 195 26 386.3
; EARLIER FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Papilloma virus
US-09-000-266-5

Query Match          90.9%; Score 50; DB 4; Length 144;
Best Local Similarity 77.8%; Pred. No. 0.085;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AQIFNKPYW 9
Db      124 AQLFNRPYW 132

RESULT 24
US-09-628-099-5
; Sequence 5, Application US/09628099
; Patent No. 6368832
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999

```

; CURRENT APPLICATION NUMBER: US/09/628,099
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Papilloma virus
US-09-628-099-5

Query Match 90.9%; Score 50; DB 4; Length 144;
Best Local Similarity 77.8%; Pred. No. 0.085;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
||:||||
Db 124 AQLFNRPYW 132

RESULT 25

US-10-056-360-5
; Sequence 5, Application US/10056360
; Patent No. 6555345
; GENERAL INFORMATION:

; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald

; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999

; CURRENT APPLICATION NUMBER: US/10/056,360
; CURRENT FILING DATE: 2000-01-23

; PRIOR APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19

; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19

; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5

; LENGTH: 144
; TYPE: PRT

; ORGANISM: Papilloma virus
US-10-056-360-5

Query Match 90.9%; Score 50; DB 4; Length 144;
Best Local Similarity 77.8%; Pred. No. 0.085;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
||:||||
Db 124 AQLFNRPYW 132

RESULT 26

US-10-056-359-5
; Sequence 5, Application US/10056359
; Patent No. 6562597
; GENERAL INFORMATION:

; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald

; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999

; CURRENT APPLICATION NUMBER: US/10/056,359
; CURRENT FILING DATE: 2002-01-23

; PRIOR APPLICATION NUMBER: US/09/628,099

; PRIOR FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Papilloma virus
US-10-056-359-5

Query Match 90.9%; Score 50; DB 4; Length 144;
Best Local Similarity 77.8%; Pred. No. 0.085;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQIFNKPYW 9
||:||||
Db 124 AQLFNRPYW 132

RESULT 27

US-08-409-122-2
; Sequence 2, Application US/08409122
; Patent No. 5820870
; GENERAL INFORMATION:

; APPLICANT: JOYCE, JAMES G.
; APPLICANT: GEORGE, HUGH A.
; APPLICANT: HOFMANN, KATHRYN J.

; APPLICANT: JANSEN, KATHRYN U.
; APPLICANT: NEEPER, MICHAEL P.

; TITLE OF INVENTION: DNA ENCODING HUMAN PAPILLOMAVIRUS TYPE 18 VACCINE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CHRISTINE E. CARTY - MERCK & CO., INC.
; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
; CITY: RAHWAY

; STATE: NJ
; COUNTRY: US

; ZIP: 07065-0907
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/409,122
; FILING DATE:

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/408,669
; FILING DATE: 22-MAR-1995

; ATTORNEY/AGENT INFORMATION:
; NAME: CARTY, CHRISTINE E

; REGISTRATION NUMBER: 36,099
; REFERENCE/DOCKET NUMBER: 19425

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-6734

; TELEFAX: 908-594-4720
; TELEX:

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 507 amino acids
; TYPE: amino acid

; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO

; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:
US-08-409-122-2

Query Match 90.9%; Score 50; DB 2; Length 507;
Best Local Similarity 77.8%; Pred. No. 0.31;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
:|||||
Db 304 SQLFNKPYW 312

RESULT 28
US-08-408-669-2
; Sequence 2, Application US/08408669
; Patent No. 5840306
; GENERAL INFORMATION:
; APPLICANT: HOFMANN, KATHRYN J.
; APPLICANT: JANSSEN, KATHRIN U.
; APPLICANT: NEPPER, MICHAEL P.
; TITLE OF INVENTION: DNA ENCODING HUMAN PAPILLOMAVIRUS TYPE 18
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESS: CHRISTINE E. CARTY - MERCK & CO., INC.
; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/408,669
; FILING DATE: 22-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CARTY, CHRISTINE E
; REGISTRATION NUMBER: 36,099
; REFERENCE/DOCKET NUMBER: 19424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-6734
; TELEFAX: 908-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 507 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-08-408-669-2

Query Match 90.9%; Score 50; DB 2; Length 507;
Best Local Similarity 77.8%; Pred. No. 0.31;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
:|||||
Db 304 SQLFNKPYW 312

RESULT 29
US-09-520-822A-3
; Sequence 3, Application US/09520822A
; Patent No. 6551597

; GENERAL INFORMATION:

; APPLICANT: Harrison, Stephen
; APPLICANT: Chen, Xiaojiang
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
; TITLE OF INVENTION: virus
; FILE REFERENCE: Harvard/Harrison 12687/1120
; CURRENT APPLICATION NUMBER: US/09/520,822A
; CURRENT FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/125208
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/148544
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Human papillomavirus type 18
US-09-520-822A-3

Query Match 90.9%; Score 50; DB 4; Length 568;
Best Local Similarity 77.8%; Pred. No. 0.34;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
:|||||
Db 365 SQLFNKPYW 373

RESULT 30
US-09-319-056B-11
; Sequence 11, Application US/09319056B
; Patent No. 6413522
; GENERAL INFORMATION:
; APPLICANT: de Villiers-zur Hausen, Ethel-Michele
; APPLICANT: zur Hausen, Harald
; APPLICANT: Laverne, Donna
; APPLICANT: Benton, Claire
; TITLE OF INVENTION: PAPILLOMA VIRUSES, PRODUCTS FOR THE
; TITLE OF INVENTION: DETECTION THEREOF AS WELL AS FOR TREATING DISEASES CAUSED BY
; FILE REFERENCE: 8484-0086-999
; CURRENT APPLICATION NUMBER: US/09/319,056B
; CURRENT FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: PCT/DE97/02659
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: DE 196 48 962.8
; PRIOR FILING DATE: 1996-11-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Papilloma virus
US-09-319-056B-11

Query Match 89.1%; Score 49; DB 4; Length 127;
Best Local Similarity 87.5%; Pred. No. 0.11;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QIFNKPYW 9
:|||||
Db 110 QLFNKPYW 117

RESULT 31
US-09-000-266-8
; Sequence 8, Application US/09000266A
; Patent No. 6322795
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald


```
; EARLIER FILING DATE: 1996-07-19
; EARLIER APPLICATION NUMBER: DE P 195 26 386.3
; EARLIER FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Papilloma virus
US-09-000-266-26

Query Match      89.1%; Score 49; DB 4; Length 141;
Best Local Similarity 77.8%; Pred. No. 0.13;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
Db 121 SQIFNRPYW 129

RESULT 36
US-09-628-099-26
; Sequence 26, Application US/09628099
; Patent No. 6368832
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/09/628,099
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: DE P 195 26 386.3
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Papilloma virus
US-09-628-099-26

Query Match      89.1%; Score 49; DB 4; Length 141;
Best Local Similarity 77.8%; Pred. No. 0.13;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQIFNKPYW 9
Db 121 SQIFNRPYW 129

RESULT 37
US-10-056-360-26
; Sequence 26, Application US/10056360
; Patent No. 655345
; GENERAL INFORMATION:
; APPLICANT: Shamanin, Vladimir
; APPLICANT: De Villiers-Zur Hausen, Ethel-Michele
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: PAPILLOMA VIRUSES, AGENTS FOR DETECTING
; TITLE OF INVENTION: THEM AND FOR TREATING DISEASES CAUSED BY SUCH VIRUSES
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/10/056,360
; CURRENT FILING DATE: 2000-01-23
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: 09/000,266
; PRIOR FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; FILE REFERENCE: 8484-0037-999
; CURRENT APPLICATION NUMBER: US/09/479,645A
; CURRENT FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: PCT/AU98/00530
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: AU P07765
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: AU P09467
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 2
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Bovine papillomavirus type 1
; FEATURE:
; OTHER INFORMATION: L1 open reading frame (wild-type)
US-09-479-645A-2

Query Match      87.3%; Score 48; DB 4; Length 495;
Best Local Similarity 87.5%; Pred. NO. 0.68;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 QIFNRPYW 9
      |||:|
Db      302 QIFNRPYW 309
```

```
RESULT 40
US-09-479-645A-4
; Sequence 4, Application US/09479645A
; Patent No. 6489141
; GENERAL INFORMATION:
; APPLICANT: FRAZER, Ian Hector
; APPLICANT: ZHOU, Jian
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY
; FILE REFERENCE: 210338.0001/1US
; CURRENT APPLICATION NUMBER: US/09/479,645A
; CURRENT FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: PCT/AU98/00530
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: AU P07765
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: AU P09467
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Bovine
; OTHER INFORMATION: Papillomavirus type 1 L1 open reading frame
; OTHER INFORMATION: (humanized)
; FEATURE:
; OTHER INFORMATION: wild-type codons replaced with synonymous codons
; OTHER INFORMATION: used at relatively high frequency by human genes
US-09-479-645A-4
```

```
Query Match      87.3%; Score 48; DB 4; Length 495;
Best Local Similarity 87.5%; Pred. NO. 0.68;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 QIFNRPYW 9
      |||:|
Db      302 QIFNRPYW 309
```

Search completed: March 3, 2004, 10:26:09
Job time : 16.5 secs